

Two searched: attached  
09/19/0185 +  
09/1492135

Access DB# \_\_\_\_\_

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: JANE ZARRA Examiner #: 77512 Date: 12/3/02  
Art Unit: 1635 Phone Number 6-5820 Serial Number: 09/910185  
Mail Box and Bldg/Room Location: 11 DC Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: A's Model of Chemicals Oversee - 3

Inventors (please provide full names): Bennett et al

Earliest Priority Filing Date: 7/18/01

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search Seq ID No: 3  
- unlimited size search  
- limit to 100 NT

Point of Contact:  
Beverly Shears  
Technical Info. Specialist  
CM1 1E05 Tel: 308-4994

Therbo

NA-5055

~~AA~~

### STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Beverly 24999</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: _____	Bibliographic _____	Dr.Link _____
Date Completed: <u>12-11-02</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: <u>5</u>	Fulltext _____	Sequence Systems _____
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: <u>20</u>	Other _____	Other (specify) <u>CGN</u>

PTO-1590 (8-01)



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2002, 01:25:16 ; Search time 8359 Seconds

(without alignments)  
17599.555 Million cell updates/sec

Title: US-09-910-185-3

Perfect score: 5055

Sequence: 1 cgatactacggggcatttt.....accctcttttaaaaaaa 5055

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing:

Maximum Match 0%

Listing first 1000 summaries

Database :  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_ov:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.6	0.6	31	AX249344	AX249344 Sequence
2	30.6	0.6	31	AX249345	AX249345 Sequence
3	29.6	0.6	78	ARI59551	ARI59551 Sequence
4	28	0.6	69	ARI59553	ARI59553 Sequence
5	28	0.6	78	ARI59552	ARI59552 Sequence
6	28	0.6	88	AX039313	AX039313 Sequence
7	28	0.6	88	AX039522	AX039522 Sequence
8	27.6	0.5	97	AF223423	AF223423 Pseudomon
9	27.4	0.5	99	ARI170415	ARI170415 Sequence
10	27.4	0.5	99	IR3148	IR3148 Sequence 10
11	27.2	0.5	98	PTU18443	PTU18443 Pan troglod
12	27	0.5	93	XU082800	XU082800 Xiphophorus
13	27	0.5	96	AX031024	AX031024 Sequence
14	27	0.5	96	AX031034	AX031034 Sequence
15	27	0.5	99	AX031011	AX031011 Sequence
16	26.6	0.5	99	AX031041	AX031041 Sequence
17	26.6	0.5	82	HSPE06C12	AL009338 H.sapiens
18	26.6	0.5	88	SSGGRP1	X16533 S. scrofa D
19	26.4	0.5	75	ARI59550	ARI59550 Sequence
20	26.4	0.5	100	HS1ASA	M13884 Human herpe
21	26.2	0.5	63	AF189453	AF189453 Homo sapi
22	26.2	0.5	99	RRCGR9	X69669 R.ratus (S
23	26.2	0.5	99	AB032281	AB032281 Tr vitru
24	26	0.5	66	AF2702	AF2702 Sequence 3
25	26	0.5	86	AX039312	AX039312 Sequence
26	26	0.5	86	AX039521	AX039521 Sequence
27	25.8	0.5	74	S79494	S79494 mex40 [Drea
28	25.8	0.5	97	AX023588	AX023588 Sequence
29	25.6	0.5	82	HCH38053	AJ388053 HyLa chry
30	25.6	0.5	84	AR021158	AR021158 Sequence
31	25.6	0.5	96	AX363260	AX363260 Sequence
32	25.4	0.5	99	AX015181	AX015181 Sequence
33	25.2	0.5	88	MMAPL2	U21945 Mus musculu
34	25.2	0.5	93	AX320853	AX320853 Sequence
35	25.2	0.5	94	AX008878	AX008878 Sequence
36	25.2	0.5	96	AL841018	AL841018 Arabidops
37	25.2	0.5	100	AF235063	AF235063 Mus muscu
38	25	0.5	25	HSR270365	AJ270365 Homo sapi
39	25	0.5	98	AX039508	AX039508 Sequence
40	25	0.5	98	AX039517	AX039517 Sequence
41	25	0.5	100	AX039509	AX039509 Sequence
42	25	0.5	100	AX039518	AX039518 Sequence
43	25	0.5	100	HUMZNP8	MB8364 Homo sapien
44	24.8	0.5	77	AX039314	AX039314 Sequence
45	24.8	0.5	77	AX039523	AX039523 Sequence
46	24.8	0.5	79	AX039315	AX039315 Sequence
47	24.8	0.5	79	AX039524	AX039524 Sequence
48	24.8	0.5	84	I13294	I13294 Sequence 7
49	24.8	0.5	86	AX039312	AX039312 Sequence
50	24.8	0.5	86	AX039521	AX039521 Sequence
51	24.8	0.5	88	AX039313	AX039313 Sequence
52	24.8	0.5	88	AX039522	AX039522 Sequence
53	24.8	0.5	92	AX039310	AX039310 Sequence
54	24.8	0.5	92	AX039519	AX039519 Sequence
55	24.8	0.5	94	AX039311	AX039311 Sequence
56	24.8	0.5	94	AX039520	AX039520 Sequence
57	24.8	0.5	98	AX039517	AX039517 Sequence
58	24.8	0.5	98	AX039518	AX039518 Sequence
59	24.8	0.5	100	AX039519	AX039519 Sequence
60	24.8	0.5	100	AX039520	AX039520 Sequence
61	24.8	0.5	100	AJ08121	AJ08121 Homo sapi
62	24.6	0.5	92	AX039310	AX039310 Sequence
63	24.6	0.5	92	AX039519	AX039519 Sequence
64	24.6	0.5	94	AX039311	AX039311 Sequence
65	24.6	0.5	94	AX039520	AX039520 Sequence

Pred. No. is the number of results predicted by chance to have a

66	24.6	0.5	100	14	AB032315	AB032315 TT virus	139	23.4	0.5	75	6	AR043032	AR043032 Sequence
67	24.6	0.5	100	14	AB032318	AB032318 TT virus	140	23.4	0.5	75	6	AR161320	AR161320 Sequence
68	24.4	0.5	72	6	AR021452	AR021452 Sequence	141	23.4	0.5	75	6	I63023	I63023 Sequence
69	24.4	0.5	72	6	AR043014	AR043014 Sequence	142	23.4	0.5	75	6	I63023	I63023 Sequence
70	24.4	0.5	72	6	I63005	I63005 Sequence	143	23.4	0.5	76	6	AR161313	AR161313 Sequence
71	24.4	0.5	72	6	I63005	I63005 Sequence	144	23.4	0.5	77	6	AR021463	AR021463 Sequence
72	24.4	0.5	81	6	AX202479	AX202479 Sequence	145	23.4	0.5	77	6	AR043025	AR043025 Sequence
73	24.4	0.5	87	6	AX384806	AX384806 Sequence	146	23.4	0.5	77	6	AX039314	AX039314 Sequence
74	24.4	0.5	90	6	AB076771	AB076771 Homo sapi	147	23.4	0.5	77	6	AX039523	AX039523 Sequence
75	24.4	0.5	92	9	AY006181	AY006181 Homo sapi	148	23.4	0.5	77	6	I63016	I63016 Sequence
76	24.4	0.5	97	6	AX388946	AX388946 Sequence	149	23.4	0.5	77	6	I63016	I63016 Sequence
77	24.4	0.5	97	10	MUSASB	M82877 Mus musculus	150	23.4	0.5	78	6	AX328773	AX328773 Sequence
78	24.2	0.5	57	6	ARI59554	ARI59554 Sequence	151	23.4	0.5	79	6	AX039315	AX039315 Sequence
79	24.2	0.5	69	6	I13295	I13295 Sequence	152	23.4	0.5	79	6	AX039524	AX039524 Sequence
80	24.2	0.5	90	6	AX326608	AX326608 Sequence	153	23.4	0.5	80	6	AR021456	AR021456 Sequence
81	24.2	0.5	90	6	AX326653	AX326653 Sequence	154	23.4	0.5	80	6	AR043018	AR043018 Sequence
82	24.2	0.5	93	9	AF172205	AF172205 Nycticebu	155	23.4	0.5	80	6	AR161306	AR161306 Sequence
83	24.2	0.5	93	9	AF172206	AF172206 Nycticebu	156	23.4	0.5	80	6	I63009	I63009 Sequence
84	24.2	0.5	93	9	AF172207	AF172207 Nycticebu	157	23.4	0.5	81	6	I63009	I63009 Sequence
85	24.2	0.5	94	6	AX088788	AX088788 Sequence	158	23.4	0.5	81	6	AR021450	AR021450 Sequence
86	24.2	0.5	96	9	AB016202	AB016202 Lemur cat	159	23.4	0.5	81	6	AR043012	AR043012 Sequence
87	24.2	0.5	99	6	ARI70414	ARI70414 Sequence	160	23.4	0.5	81	6	AR043012	AR043012 Sequence
88	24.2	0.5	99	6	I73147	I73147 Sequence	161	23.4	0.5	81	6	I63003	I63003 Sequence
89	24.2	0.5	100	6	AX088810	AX088810 Sequence	162	23.4	0.5	81	6	I63003	I63003 Sequence
90	24	0.5	24	9	HSAR270366	HSAR270366 Homo sapi	163	23.4	0.5	84	6	AR021460	AR021460 Sequence
91	24	0.5	60	3	AF320180	AF320180 Drosophila	164	23.4	0.5	84	6	AR021465	AR021465 Sequence
92	24	0.5	69	6	AR021449	AR021449 Sequence	165	23.4	0.5	84	6	AR043022	AR043022 Sequence
93	24	0.5	69	6	AR043011	AR043011 Sequence	166	23.4	0.5	84	6	AR043027	AR043027 Sequence
94	24	0.5	69	6	I63002	I63002 Sequence	167	23.4	0.5	84	6	AR081482	AR081482 Sequence
95	24	0.5	69	6	I63002	I63002 Sequence	168	23.4	0.5	84	6	AR161310	AR161310 Sequence
96	24	0.5	69	6	I63002	I63002 Sequence	169	23.4	0.5	84	6	AR161315	AR161315 Sequence
97	24	0.5	72	6	ARI61302	ARI61302 Sequence	170	23.4	0.5	84	6	I63013	I63013 Sequence
98	24	0.5	84	6	AR021157	AR021157 Sequence	171	23.4	0.5	84	6	I63018	I63018 Sequence
99	24	0.5	84	6	AR021159	AR021159 Sequence	172	23.4	0.5	84	6	I63018	I63018 Sequence
100	24	0.5	87	6	AX6803	AX6803 Sequence	173	23.4	0.5	84	6	I63018	I63018 Sequence
101	24	0.5	87	6	AR075881	AR075881 Sequence	174	23.4	0.5	84	6	I63018	I63018 Sequence
102	24	0.5	87	6	AR097670	AR097670 Sequence	175	23.4	0.5	85	5	CHKOI110	CHKOI110 Sequence
103	24	0.5	91	6	AX052944	AX052944 Sequence	176	23.4	0.5	85	5	AR021468	AR021468 Sequence
104	24	0.5	97	6	AX455930	AX455930 Sequence	177	23.4	0.5	85	6	AR043030	AR043030 Sequence
105	24	0.5	98	6	AX320854	AX320854 Sequence	178	23.4	0.5	85	6	AR161318	AR161318 Sequence
106	24	0.5	100	6	AX3853	AX3853 Sequence	179	23.4	0.5	85	6	I63021	I63021 Sequence
107	24	0.5	100	6	AX3856	AX3856 Sequence	180	23.4	0.5	85	6	I63021	I63021 Sequence
108	24	0.5	100	6	I17430	I17430 Sequence	181	23.4	0.5	87	6	AR021469	AR021469 Sequence
109	24	0.5	100	6	I17430	I17430 Sequence	182	23.4	0.5	87	6	AR043031	AR043031 Sequence
110	23.8	0.5	45	6	ARI94142	ARI94142 Sequence	183	23.4	0.5	87	6	AR161319	AR161319 Sequence
111	23.8	0.5	60	6	HUMSAU3A51	HUMSAU3A51 Sequence	184	23.4	0.5	87	6	I63022	I63022 Sequence
112	23.8	0.5	61	6	ARI18190	ARI18190 Sequence	185	23.4	0.5	87	6	I63022	I63022 Sequence
113	23.8	0.5	82	6	AX328774	AX328774 Sequence	186	23.4	0.5	90	6	AR021451	AR021451 Sequence
114	23.8	0.5	83	9	S72917	S72917 coxsacki	187	23.4	0.5	90	6	AR043013	AR043013 Sequence
115	23.8	0.5	90	6	AB9632	AB9632 Sequence	188	23.4	0.5	90	6	AX088807	AX088807 Sequence
116	23.8	0.5	90	6	ARI76127	ARI76127 Sequence	189	23.4	0.5	90	6	I63004	I63004 Sequence
117	23.8	0.5	91	6	AF330163	AF330163 Canis fam	190	23.4	0.5	90	6	I63004	I63004 Sequence
118	23.8	0.5	91	6	I07154	I07154 Sequence	191	23.4	0.5	93	6	ARI61301	ARI61301 Sequence
119	23.8	0.5	91	6	I07202	I07202 Sequence	192	23.4	0.5	94	6	AX480654	AX480654 Sequence
120	23.8	0.5	93	9	HSU15371	HSU15371 Human andro	193	23.4	0.5	96	6	AX480661	AX480661 Sequence
121	23.8	0.5	94	6	AX023601	AX023601 Sequence	194	23.4	0.5	96	6	I58482	I58482 Sequence
122	23.8	0.5	94	6	AX068890	AX068890 Sequence	195	23.4	0.5	97	4	BTBKH2	BTBKH2 Sequence
123	23.8	0.5	94	11	HUMSMX2453	HUMSMX2453 Sequence	196	23.4	0.5	98	6	ARI25349	ARI25349 Sequence
124	23.8	0.5	96	5	AF459839	AF459839 Gallus ga	197	23.4	0.5	99	3	DROHIGB04	DROHIGB04 Sequence
125	23.8	0.5	96	6	ARI06481	ARI06481 Sequence	198	23.4	0.5	100	6	AX404584	AX404584 Sequence
126	23.8	0.5	96	6	AF395210	AF395210 Homo sapi	199	23.2	0.5	51	6	AX150270	AX150270 Sequence
127	23.6	0.5	79	6	ARI76191	ARI76191 Sequence	200	23.2	0.5	51	6	AX162355	AX162355 Sequence
128	23.6	0.5	79	6	E22906	E22906 Artificial	201	23.2	0.5	51	6	AX162356	AX162356 Sequence
129	23.6	0.5	80	6	AR201912	AR201912 Sequence	202	23.2	0.5	81	6	S76506	S76506 Sequence
130	23.6	0.5	82	5	FSAXHXPME	FSAXHXPME Sequence	203	23.2	0.5	81	6	AR030404	AR030404 Sequence
131	23.6	0.5	90	5	XHU82793	XHU82793 Xiphophorus	204	23.2	0.5	81	6	I40762	I40762 Sequence
132	23.6	0.5	90	5	XHU82799	XHU82799 Xiphophorus	205	23.2	0.5	83	6	AR051740	AR051740 Sequence
133	23.6	0.5	99	6	A08898	A08898 H. sapiens (	206	23.2	0.5	83	6	AR200063	AR200063 Sequence
134	23.6	0.5	100	5	XLRLP26	XLRLP26 Xenopus lae	207	23.2	0.5	83	6	BD002010	BD002010 Sequence
135	23.6	0.5	100	11	HSAR10741	HSAR10741 Homo sapi	208	23.2	0.5	83	6	I08745	I08745 Sequence
136	23.4	0.5	60	9	SS3635	SS3635 APOA-IV (Ap	209	23.2	0.5	83	6	I49686	I49686 Sequence
137	23.4	0.5	71	6	AR066445	AR066445 Sequence	210	23.2	0.5	90	6	AX326609	AX326609 Sequence
138	23.4	0.5	75	6	AR021470	AR021470 Sequence	211	23.2	0.5	90	6	AX326666	AX326666 Sequence



c 212	23.2	0.5	91	6	AR205306	AR205306 Sequence	c 285	22.6	0.4	59	6	E13436	E13436 Synthetic o
c 213	23.2	0.5	93	6	AX167441	AX167441 Sequence	c 286	22.6	0.4	73	6	AX342432	AX342432 Sequence
c 214	23.2	0.5	97	6	AR068361	AR068361 Sequence	c 287	22.6	0.4	75	6	I49876	I49876 Sequence
c 215	23.2	0.5	99	5	S82829	S82829 FGFI=IltroD	c 288	22.6	0.4	76	6	H0MGRP2	H0MGRP2
c 216	23	0.5	64	5	PERN28D8	Z18739 Perca fluvl	c 289	22.6	0.4	79	6	AX057967	AX057967 Sequence
c 217	23	0.5	68	6	AR007249	AR007249 Sequence	c 290	22.6	0.4	79	6	AX057968	AX057968 Sequence
c 218	23	0.5	68	6	AR062453	AR062453 Sequence	c 291	22.6	0.4	82	9	HSPHKG07	HSPHKG07
c 219	23	0.5	68	6	AR170035	AR170035 Sequence	c 292	22.6	0.4	86	8	OSA307915	OSA307915
c 220	23	0.5	68	6	AR170462	AR170462 Sequence	c 293	22.6	0.4	86	10	MUSRP2A	MUSRP2A
c 221	23	0.5	68	6	I27739	I27739 Sequence	c 294	22.6	0.4	90	11	HSAR248353	HSAR248353
c 222	23	0.5	78	6	AX278071	AX278071 Sequence	c 295	22.6	0.4	90	14	HSILPA	HSILPA
c 223	23	0.5	81	14	AY079300	AY079300 Hepatitis	c 296	22.6	0.4	93	6	AR051773	AR051773 Sequence
c 224	23	0.5	81	14	AY079301	AY079301 Hepatitis	c 297	22.6	0.4	93	6	AR051776	AR051776 Sequence
c 225	23	0.5	82	6	A36805	A36805 Sequence	c 298	22.6	0.4	93	9	AF172208	AF172208 Eulemur m
c 226	23	0.5	82	6	AR075883	AR075883 Sequence	c 299	22.6	0.4	93	9	AF172209	AF172209 Perodicti
c 227	23	0.5	82	6	AR097672	AR097672 Sequence	c 300	22.6	0.4	93	9	HUMGSTF	HUMGSTF
c 228	23	0.5	82	10	AF265865	AF265865 Mus muscu	c 301	22.6	0.4	96	6	I34361	I34361
c 229	23	0.5	83	6	AX080696	AX080696 Sequence	c 302	22.6	0.4	96	6	I34373	I34373
c 230	23	0.5	85	6	AX099459	AX099459 Sequence	c 303	22.6	0.4	97	10	AF417933S1	AF417933 Mus muscu
c 231	23	0.5	89	6	AR074468	AR074468 Sequence	c 304	22.6	0.4	98	11	G33091	G33091 Eb84975 Hum
c 232	23	0.5	89	6	AR081148	AR081148 Sequence	c 305	22.6	0.4	99	6	AX068170	AX068170 Sequence
c 233	23	0.5	89	6	AR085345	AR085345 Sequence	c 306	22.6	0.4	100	6	I34360	I34360 Sequence
c 234	23	0.5	89	6	AR088093	AR088093 Sequence	c 307	22.6	0.4	100	6	I34372	I34372 Sequence
c 235	23	0.5	89	6	AR104252	AR104252 Sequence	c 308	22.4	0.4	54	6	AX008237	AX008237 Sequence
c 236	23	0.5	89	6	AR143516	AR143516 Sequence	c 309	22.4	0.4	60	3	AF320168	AF320168 Drosophila
c 237	23	0.5	89	6	AR171420	AR171420 Sequence	c 310	22.4	0.4	65	6	AX484855	AX484855 Sequence
c 238	23	0.5	89	6	AR171591	AR171591 Sequence	c 311	22.4	0.4	65	6	AX484854	AX484854 Sequence
c 239	23	0.5	90	6	AX088777	AX088777 Sequence	c 312	22.4	0.4	65	6	AX486057	AX486057 Sequence
c 240	23	0.5	91	6	HS033802	HS033802 Human immat	c 313	22.4	0.4	67	4	S67757	S67757 prostaeycli
c 241	23	0.5	91	11	HS03381BP	X97939 H. sapiens S	c 314	22.4	0.4	69	6	AR161299	AR161299 Sequence
c 242	23	0.5	93	6	AX440000	AX440000 Sequence	c 315	22.4	0.4	72	9	I13713	I13713 Sequence
c 243	23	0.5	93	6	AX440056	AX440056 Sequence	c 316	22.4	0.4	72	9	HS081101	HS081101
c 244	23	0.5	96	6	A62763	A62763 Sequence	c 317	22.4	0.4	77	9	S64415	S64415
c 245	23	0.5	96	6	I73508	I73508 Sequence	c 318	22.4	0.4	80	1	HYRNFS	HYRNFS
c 246	23	0.5	96	10	Y033555S12	Y0335556 Mus muscu	c 319	22.4	0.4	81	6	AX022983	AX022983 Sequence
c 247	23	0.5	96	14	S51471S1	S51471 gag/jund-Ga	c 320	22.4	0.4	89	6	AX046264	AX046264 Sequence
c 248	23	0.5	96	14	AX080697	AX080697 Sequence	c 321	22.4	0.4	89	6	AX052945	AX052945 Sequence
c 249	23	0.5	99	6	I14151	I14151 Sequence	c 322	22.4	0.4	90	5	AF035412	AF035412 Xiphophor
c 250	23	0.5	99	6	I14152	I14152 Sequence	c 323	22.4	0.4	90	5	XX0882796	XX0882796 Xiphophor
c 251	23	0.5	99	9	AY006143	AY006143 Homo sapi	c 324	22.4	0.4	91	9	HSANJ10679	HSANJ10679 Homo sapi
c 252	23	0.5	99	10	MM0011105	AJ011105 Mus muscu	c 325	22.4	0.4	92	3	BME1BR3	BME1BR3
c 253	23	0.5	100	9	AJ8146	AJ008146 Homo sapi	c 326	22.4	0.4	93	10	RNU08886	RNU08886
c 254	23	0.5	100	9	HSAT05	X03457 Human satel	c 327	22.4	0.4	94	1	STU06132	STU06132
c 255	23	0.5	100	14	AB032317	AB032317 TT virus	c 328	22.4	0.4	94	5	MTABAI8SB	MTABAI8SB
c 256	22.8	0.5	50	6	AR032654	AR032654 Sequence	c 329	22.4	0.4	94	5	MTATRI8SB	MTATRI8SB
c 257	22.8	0.5	50	6	AR032654	AR032654 Sequence	c 330	22.4	0.4	94	5	MTATRI8SB	MTATRI8SB
c 258	22.8	0.5	50	6	AR209308	AR209308 Sequence	c 331	22.4	0.4	94	5	MTATRI8SB	MTATRI8SB
c 259	22.8	0.5	50	6	AR209318	AR209318 Sequence	c 332	22.4	0.4	94	5	MTATRI8SB	MTATRI8SB
c 260	22.8	0.5	50	6	I29384	I29384 Sequence	c 333	22.4	0.4	94	5	MTATRI8SB	MTATRI8SB
c 261	22.8	0.5	50	6	I29394	I29394 Sequence	c 334	22.4	0.4	95	6	A36801	A36801
c 262	22.8	0.5	50	6	I91058	I91058 Sequence	c 335	22.4	0.4	95	6	AR075879	AR075879 Sequence
c 263	22.8	0.5	57	10	I91068	I91068 Sequence	c 336	22.4	0.4	95	6	AR097668	AR097668 Sequence
c 264	22.8	0.5	57	10	RNU59456	U59456 Rattus norv	c 337	22.4	0.4	96	6	AR08909	AR08909
c 265	22.8	0.5	65	6	AX485349	AX485349 Sequence	c 338	22.4	0.4	96	6	I73509	I73509
c 266	22.8	0.5	71	5	LRN28D8	Z18735 Limphorus	c 339	22.4	0.4	98	6	BD011910	BD011910
c 267	22.8	0.5	75	5	XM082797	XM082797 Sequence	c 340	22.4	0.4	98	6	BD011916	BD011916 Therapeut
c 268	22.8	0.5	75	5	XM082802	XM082802 Xiphophorus	c 341	22.4	0.4	98	6	BD012027	BD012027 Therapeut
c 269	22.8	0.5	76	6	AR003787	AR003787 Sequence	c 342	22.4	0.4	98	6	BD012914	BD012914
c 270	22.8	0.5	76	6	AR010123	AR010123 Sequence	c 343	22.4	0.4	98	6	E23310	E23310
c 271	22.8	0.5	76	6	AR055329	AR055329 Sequence	c 344	22.4	0.4	98	6	E27079	E27079
c 272	22.8	0.5	76	6	AR055329	AR055329 Sequence	c 345	22.4	0.4	98	6	BD004446	BD004446
c 273	22.8	0.5	76	6	AR141258	AR141258 Sequence	c 346	22.4	0.4	98	23	BD004446	BD004446 Therapeut
c 274	22.8	0.5	76	6	AR141495	AR141495 Sequence	c 347	22.4	0.4	98	23	BD004446	BD004446 Therapeut
c 275	22.8	0.5	76	6	I11969	I11969 Sequence	c 348	22.4	0.4	98	23	BD008340	BD008340
c 276	22.8	0.5	76	6	I40540	I40540 Sequence	c 349	22.4	0.4	98	23	BD008340	BD008340
c 277	22.8	0.5	83	9	S59798S13	S59832 PKLR-L-type	c 350	22.4	0.4	99	6	AX023606	AX023606
c 278	22.8	0.5	85	12	SYNSLMTIRE	M12614 ChimERIC D.	c 351	22.2	0.4	51	6	AX160861	AX160861 Sequence
c 279	22.8	0.5	86	6	AR042803	AR042803 Sequence	c 352	22.2	0.4	51	6	AX024250	AX024250 Sequence
c 280	22.8	0.5	86	6	AX039266	AX039266 Sequence	c 353	22.2	0.4	56	6	AX381233	AX381233 Sequence
c 281	22.8	0.5	99	6	AX080700	AX080700 Sequence	c 354	22.2	0.4	57	6	AR014088	AR014088 Sequence
c 282	22.8	0.5	99	9	AB016201	AB016201 Sequence	c 355	22.2	0.4	57	6	AR117057	AR117057 Sequence
c 283	22.8	0.5	99	10	RN078142	U78142 Rattus norv	c 356	22.2	0.4	57	6	AR117058	AR117058 Sequence
c 284	22.6	0.4	59	6	E13421	E13421 Synthetic o	c 357	22.2	0.4	57	6	I20517	I20517 Sequence

C 358	22.2	0.4	57	6	120518	Sequence 52	431	22	0.4	80	6	E29111	Kit for dia
C 359	22.2	0.4	58	6	AX358700	Sequence	C 432	22	0.4	81	6	I86928	Sequence 17
C 360	22.2	0.4	58	6	AX358701	Sequence	C 433	22	0.4	81	6	HUM5F08M3	Human Hepc2
C 361	22.2	0.4	58	6	AX358707	Sequence	C 434	22	0.4	81	9	HSU5088	Human ISOLA
C 362	22.2	0.4	58	6	AX358708	Sequence	C 435	22	0.4	82	5	AF044801	AF044801 Lampetra
C 363	22.2	0.4	58	6	AX358713	Sequence	C 436	22	0.4	84	6	A46818	A46818 Sequence 15
C 364	22.2	0.4	58	6	AX358714	Sequence	C 437	22	0.4	86	9	F185592S08	F185592S08
C 365	22.2	0.4	60	6	AX382509	Sequence	C 438	22	0.4	88	14	APH0V15E	APH0V15E
C 366	22.2	0.4	60	6	AX382518	Sequence	C 439	22	0.4	89	6	A39999	A39999
C 367	22.2	0.4	64	6	AF080801	Sequence 11	C 440	22	0.4	89	6	A46511	A46511
C 368	22.2	0.4	64	9	AF328551	Homo sapi	C 441	22	0.4	90	5	XH082794	XH082794
C 369	22.2	0.4	66	6	AX080586	Sequence	C 442	22	0.4	90	5	XM082798	XM082798
C 370	22.2	0.4	67	6	AR072415	Sequence	C 443	22	0.4	90	5	XM082801	XM082801
C 371	22.2	0.4	67	6	I06253	Sequence 11	C 444	22	0.4	90	5	XM082805	XM082805
C 372	22.2	0.4	67	6	I26526	Sequence 21	C 445	22	0.4	90	5	XM082806	XM082806
C 373	22.2	0.4	69	6	AF039636	Ceratitis	C 446	22	0.4	91	3	DDIACCT18A	DDIACCT18A
C 374	22.2	0.4	69	6	AX179510	Sequence	C 447	22	0.4	92	6	AR203063	AR203063
C 375	22.2	0.4	71	5	ACRN28D8	Amla calva	C 448	22	0.4	92	9	HSACG134	HSACG134
C 376	22.2	0.4	72	6	I49879	Anguilla an	C 449	22	0.4	93	9	HSAR294495	HSAR294495
C 377	22.2	0.4	73	5	AAARN28D8	Anguilla an	C 450	22	0.4	93	9	HSAR294495	HSAR294495
C 378	22.2	0.4	74	9	HUMDXSTR	Human DNA t	C 451	22	0.4	93	9	HSAR294496	HSAR294496
C 379	22.2	0.4	78	6	AX328771	Sequence	C 452	22	0.4	93	9	HSAR294497	HSAR294497
C 380	22.2	0.4	78	10	AF096391	Mus muscu	C 453	22	0.4	95	6	AR152238	AR152238
C 381	22.2	0.4	81	14	AF166728	Hepatitis	C 454	22	0.4	96	6	AR152238	AR152238
C 382	22.2	0.4	81	14	AF166734	Hepatitis	C 455	22	0.4	96	6	AR152238	AR152238
C 383	22.2	0.4	81	14	AF166737	Hepatitis	C 456	22	0.4	99	3	DR0G28V	DR0G28V
C 384	22.2	0.4	81	14	AF166744	Hepatitis	C 457	22	0.4	100	9	AY006225	AY006225
C 385	22.2	0.4	84	6	A35498	Synthetic 1	C 458	22	0.4	100	9	AY006326	AY006326
C 386	22.2	0.4	90	6	AR148129	Sequence	C 459	22	0.4	100	9	AY006328	AY006328
C 387	22.2	0.4	90	6	AX237216	Sequence	C 460	22	0.4	100	9	HS45BAF	HS45BAF
C 388	22.2	0.4	90	6	E12933	Sequence	C 461	22	0.4	100	9	HS45BAF	HS45BAF
C 389	22.2	0.4	90	6	E12933	Sequence	C 462	22	0.4	100	10	RNO2307479	RNO2307479
C 390	22.2	0.4	91	6	AX366434	Sequence	C 463	22	0.4	100	11	HS4310741	HS4310741
C 391	22.2	0.4	92	3	DEPROBO	Sequence 10	C 464	22	0.4	100	11	I32114	I32114
C 392	22.2	0.4	92	6	I72525	Sequence 10	C 465	22	0.4	100	11	I32114	I32114
C 393	22.2	0.4	93	8	BMU35417	Beauveria a	C 466	22	0.4	100	11	I32114	I32114
C 394	22.2	0.4	93	8	EA28SD1	Beauveria a	C 467	22	0.4	100	11	I32114	I32114
C 395	22.2	0.4	93	9	AB044123	Homo sapi	C 468	22	0.4	100	11	I32114	I32114
C 396	22.2	0.4	94	5	OM306956	Odontoph	C 469	22	0.4	100	11	I32114	I32114
C 397	22.2	0.4	95	14	E11248278	Sequence 20	C 470	22	0.4	100	11	I32114	I32114
C 398	22.2	0.4	96	5	AF459837	Gallus ga	C 471	22	0.4	100	11	I32114	I32114
C 399	22.2	0.4	96	5	AF459838	Gallus ga	C 472	22	0.4	100	11	I32114	I32114
C 400	22.2	0.4	96	5	AF395211	Homo sapi	C 473	22	0.4	100	11	I32114	I32114
C 401	22.2	0.4	96	9	AX287661	Sequence	C 474	22	0.4	100	11	I32114	I32114
C 402	22.2	0.4	99	6	A74519	Sequence 20	C 475	22	0.4	100	11	I32114	I32114
C 403	22.2	0.4	99	6	A77498	Sequence 20	C 476	22	0.4	100	11	I32114	I32114
C 404	22.2	0.4	99	10	AF007843	Mus muscu	C 477	22	0.4	100	11	I32114	I32114
C 405	22.2	0.4	100	6	AX020436	Sequence	C 478	22	0.4	100	11	I32114	I32114
C 406	22.2	0.4	49	6	AS0141	Sequence 6	C 479	22	0.4	100	11	I32114	I32114
C 407	22.2	0.4	51	6	AX162504	Sequence	C 480	22	0.4	100	11	I32114	I32114
C 408	22.2	0.4	51	6	AX162874	Sequence	C 481	22	0.4	100	11	I32114	I32114
C 409	22.2	0.4	51	6	AX204235	Sequence	C 482	22	0.4	100	11	I32114	I32114
C 410	22.2	0.4	51	6	ES9415	Signal pept	C 483	22	0.4	100	11	I32114	I32114
C 411	22.2	0.4	55	6	I13710	Sequence 16	C 484	22	0.4	100	11	I32114	I32114
C 412	22.2	0.4	57	6	AX012288	Sequence	C 485	22	0.4	100	11	I32114	I32114
C 413	22.2	0.4	61	6	AX381803	Sequence	C 486	22	0.4	100	11	I32114	I32114
C 414	22.2	0.4	65	6	AX483346	Sequence	C 487	22	0.4	100	11	I32114	I32114
C 415	22.2	0.4	66	6	AR007246	Sequence	C 488	22	0.4	100	11	I32114	I32114
C 416	22.2	0.4	66	6	AR062450	Sequence	C 489	22	0.4	100	11	I32114	I32114
C 417	22.2	0.4	66	6	AR170032	Sequence	C 490	22	0.4	100	11	I32114	I32114
C 418	22.2	0.4	66	6	AR170459	Sequence	C 491	22	0.4	100	11	I32114	I32114
C 419	22.2	0.4	66	6	AR170459	Sequence	C 492	22	0.4	100	11	I32114	I32114
C 420	22.2	0.4	67	6	AR072442	Sequence 23	C 493	22	0.4	100	11	I32114	I32114
C 421	22.2	0.4	67	6	AR072442	Sequence 23	C 494	22	0.4	100	11	I32114	I32114
C 422	22.2	0.4	68	6	AR161850	Sequence 24	C 495	22	0.4	100	11	I32114	I32114
C 423	22.2	0.4	74	6	AX1918	Sequence	C 496	22	0.4	100	11	I32114	I32114
C 424	22.2	0.4	74	6	AX1918	Sequence	C 497	22	0.4	100	11	I32114	I32114
C 425	22.2	0.4	75	5	XM082797	Xiphophorus	C 498	22	0.4	100	11	I32114	I32114
C 426	22.2	0.4	75	5	XM082804	Xiphophorus	C 499	22	0.4	100	11	I32114	I32114
C 427	22.2	0.4	75	5	XM082804	Xiphophorus	C 500	22	0.4	100	11	I32114	I32114
C 428	22.2	0.4	78	6	AX1917	nucleotide	C 501	22	0.4	100	11	I32114	I32114
C 429	22.2	0.4	78	6	AX1917	nucleotide	C 502	22	0.4	100	11	I32114	I32114
C 430	22.2	0.4	80	6	AB7762	Sequence 56	C 503	22	0.4	100	11	I32114	I32114

504	21.8	0.4	100	11	G32918	G32918 A009W25 Hum	577	21.4	0.4	84	9	HUMENL1	M14059 Human liver
505	21.8	0.4	100	11	HS309128	AJ309128 Homo sapi	c 578	21.4	0.4	85	6	AR030723	AR030723 Sequence
c 506	21.6	0.4	43	3	AAR2	X15295 Anopheles a	579	21.4	0.4	87	6	A36803	A36803 Sequence 22
c 507	21.6	0.4	47	6	AX195036	AX195036 Sequence	580	21.4	0.4	87	6	AR075881	AR075881 Sequence
c 508	21.6	0.4	52	6	AX456490	AX456490 Sequence	581	21.4	0.4	87	6	AR097670	AR097670 Sequence
c 509	21.6	0.4	54	6	I13484	I13484 Sequence 18	582	21.4	0.4	87	13	M23845	M23845 Figure 3. N
c 510	21.6	0.4	57	7	pp7CIN2	M12222 Bacterioph	583	21.4	0.4	88	9	S79269	S79269 Duffly [huma
c 511	21.6	0.4	59	6	AR172733	AR172733 Sequence	c 584	21.4	0.4	88	11	AF021107	AF021107 Homo sapi
c 512	21.6	0.4	59	6	AR178677	AR178677 Sequence	c 585	21.4	0.4	89	6	AX068163	AX068163 Sequence
c 513	21.6	0.4	59	6	AR181248	AR181248 Sequence	c 586	21.4	0.4	89	9	HSYAP0G11	HSYAP0G11 Homo sapi
c 514	21.6	0.4	59	6	AR199929	AR199929 Sequence	c 587	21.4	0.4	90	6	AX052903	AX052903 Sequence
c 515	21.6	0.4	64	6	A60793	A60793 Sequence 10	c 588	21.4	0.4	90	8	AF176407	AF176407 Quercus f
c 516	21.6	0.4	70	6	AR174789	AR174789 Sequence	c 589	21.4	0.4	90	8	HSCELT	HSCELT Human T lymph
c 517	21.6	0.4	72	6	AX366537	AX366537 Sequence	c 590	21.4	0.4	91	8	AF176404	AF176404 Quercus f
c 518	21.6	0.4	72	10	MUSCIA17	K03035 Mouse alpha	c 591	21.4	0.4	92	6	AR014058	AR014058 Sequence
c 519	21.6	0.4	73	6	E33309	E33309 Dsba/Dsba/D	c 592	21.4	0.4	92	6	I22008	I22008 Sequence 94
c 520	21.6	0.4	74	10	AF265986	AF265986 Mus muscu	c 593	21.4	0.4	92	14	HTU012104	HTU012104 Human T-cell
c 521	21.6	0.4	76	6	AX174956	AX174956 Sequence	c 594	21.4	0.4	94	6	AR140871	AR140871 Sequence
c 522	21.6	0.4	77	6	AX057282	AX057282 Sequence	c 595	21.4	0.4	94	6	AR150821	AR150821 Sequence
c 523	21.6	0.4	77	6	HSU29117	U29117 Human lipom	c 596	21.4	0.4	94	6	AX023633	AX023633 Sequence
c 524	21.6	0.4	79	8	CBE345109	AJ345109 Carpinus	c 597	21.4	0.4	94	6	AX182150	AX182150 Sequence
c 525	21.6	0.4	79	12	SYNML3PLK	K00611 plasmid pap	c 598	21.4	0.4	94	6	BD003074	BD003074 Polynucle
c 526	21.6	0.4	81	6	AR119865	AR119865 Sequence	c 599	21.4	0.4	94	6	I65699	I65699 Sequence 59
c 527	21.6	0.4	83	6	AX029378	AX029378 Sequence	c 600	21.4	0.4	94	6	I67931	I67931 Sequence 59
c 528	21.6	0.4	85	3	AY075043	AY075043 Polsonia	c 601	21.4	0.4	94	6	I90152	I90152 Sequence 59
c 529	21.6	0.4	87	6	AX278072	AX278072 Sequence	c 602	21.4	0.4	96	8	HVU234896	HVU234896 Hordeum v
c 530	21.6	0.4	88	6	AX146316	AX146316 Sequence	c 603	21.4	0.4	96	8	AF395211	AF395211 Homo sapi
c 531	21.6	0.4	88	8	S77500	S77500 Oxi3-cytoch	c 604	21.4	0.4	96	9	AF395792	AF395792 Homo sapi
c 532	21.6	0.4	90	5	AF035412	AF035412 Xiphophor	c 605	21.4	0.4	96	10	RATPMCA3A2	RATPMCA3A2
c 533	21.6	0.4	90	6	AX435642	AX435642 Sequence	c 606	21.4	0.4	97	6	A36806	A36806 Rat plasma
c 534	21.6	0.4	90	5	HSU5163	U5163 Human Isola	c 607	21.4	0.4	97	6	AR075884	AR075884 Sequence
c 535	21.6	0.4	90	14	AF106090	AF106090 Hepatitis	c 608	21.4	0.4	97	6	AR097673	AR097673 Sequence
c 536	21.6	0.4	90	14	AF106093	AF106093 Hepatitis	c 609	21.4	0.4	98	8	S66605	S66605 fclone t33.
c 537	21.6	0.4	90	14	AF106101	AF106101 Hepatitis	c 610	21.4	0.4	99	9	HSITGAD05	HSITGAD05 Human beta-
c 538	21.6	0.4	90	14	AF106103	AF106103 Hepatitis	c 611	21.4	0.4	100	3	AG2H117	AG2H117 A.gambiae s
c 539	21.6	0.4	93	4	FCMYOS03	AF001618 Fells cat	c 612	21.4	0.4	100	6	AR119732	AR119732 Sequence
c 540	21.6	0.4	93	6	A28234	A28234 DNA cassette	c 613	21.4	0.4	100	6	AR129862	AR129862 Sequence
c 541	21.6	0.4	93	6	A28235	A28235 Homo sapien	c 614	21.4	0.4	100	9	F185592S15	F185592S15 Homo sapi
c 542	21.6	0.4	93	6	D78279S16	D78294 Homo sapien	c 615	21.4	0.4	100	9	HUMHTIS03	HUMHTIS03 Homo sapien
c 543	21.6	0.4	94	6	A60866	A60866 Sequence 17	c 616	21.4	0.4	100	10	AR010314	AR010314 Mus muscu
c 544	21.6	0.4	95	1	AI16S4	X67456 A. israeli	c 617	21.2	0.4	30	6	E30088	E30088 Human BMP-7
c 545	21.6	0.4	95	6	AX119975	AX119975 Sequence	c 618	21.2	0.4	51	6	AX157484	AX157484 Sequence
c 546	21.6	0.4	95	6	AX138024	AX138024 Sequence	c 619	21.2	0.4	51	6	AX163065	AX163065 Sequence
c 547	21.6	0.4	96	11	G31304	G31304 sy8991-19	c 620	21.2	0.4	51	6	AX163066	AX163066 Sequence
c 548	21.6	0.4	97	6	AX339342	AX339342 Sequence	c 621	21.2	0.4	51	6	AX189852	AX189852 Sequence
c 549	21.6	0.4	97	9	HSU2MENH	X14775 Human DNA f	c 622	21.2	0.4	51	10	AF005572	AF005572 Mus muscu
c 550	21.6	0.4	98	9	HSTGMA2	X12620 Human unpro	c 623	21.2	0.4	54	6	A62703	A62703 Sequence 4
c 551	21.6	0.4	98	14	AB032277	AB032277 TT virus	c 624	21.2	0.4	55	9	S76369	S76369 PML-RARA fu
c 552	21.6	0.4	99	6	AX023653	AX023653 Sequence	c 625	21.2	0.4	60	6	AR201911	AR201911 Sequence
c 553	21.6	0.4	99	11	G71249	G71249 721694031FM	c 626	21.2	0.4	60	9	S66488	S66488 NEFH-neurof
c 554	21.4	0.4	49	6	AX162876	AX162876 Sequence	c 627	21.2	0.4	62	6	I06254	I06254 Sequence 12
c 555	21.4	0.4	56	6	AX193988	AX193988 Sequence	c 628	21.2	0.4	63	6	AX108115	AX108115 Sequence
c 556	21.4	0.4	56	6	AX139459	AX139459 Sequence	c 629	21.2	0.4	64	6	AR073784	AR073784 Sequence
c 557	21.4	0.4	62	6	AR014635	AR014635 Sequence	c 630	21.2	0.4	64	6	AR208280	AR208280 Sequence
c 558	21.4	0.4	62	6	AR078706	AR078706 Sequence	c 631	21.2	0.4	65	6	AX484900	AX484900 Sequence
c 559	21.4	0.4	62	6	BD010427	BD010427 Chimeric	c 632	21.2	0.4	65	6	AX485140	AX485140 Sequence
c 560	21.4	0.4	62	6	I07082	I07082 Sequence 8	c 633	21.2	0.4	65	6	AX485901	AX485901 Sequence
c 561	21.4	0.4	62	6	I26767	I26767 Sequence 15	c 634	21.2	0.4	66	6	I15666	I15666 Sequence 3
c 562	21.4	0.4	62	6	I83775	I83775 Sequence 15	c 635	21.2	0.4	69	6	I36671	I36671 Sequence 3
c 563	21.4	0.4	65	6	AX485948	AX485948 Sequence	c 636	21.2	0.4	71	6	AX047839	AX047839 Sequence
c 564	21.4	0.4	66	8	AF027108	AF027108 Ustilago	c 637	21.2	0.4	72	6	AX150205	AX150205 Sequence
c 565	21.4	0.4	73	5	SCRN28D8	Z18748 Scyllorhinu	c 638	21.2	0.4	72	8	MILKTRND	MILKTRND Yeast mitoc
c 566	21.4	0.4	75	6	SCRN2603	AR135603 Sequence	c 639	21.2	0.4	72	8	MILKTRND	MILKTRND Yeast mitoc
c 567	21.4	0.4	75	6	I32935	I32935 Sequence 5	c 640	21.2	0.4	73	6	AR195065	AR195065 Sequence
c 568	21.4	0.4	75	6	I32936	I32936 Sequence 6	c 641	21.2	0.4	73	6	AR212274	AR212274 Sequence
c 569	21.4	0.4	76	3	PFU06449	U06449 Plasmodium	c 642	21.2	0.4	73	6	AX277693	AX277693 Sequence
c 570	21.4	0.4	78	6	HSU91165	U91165 Homo sapien	c 643	21.2	0.4	73	9	F310192S01	F310192S01 Homo sapi
c 571	21.4	0.4	79	6	AR062571	AR062571 Sequence	c 644	21.2	0.4	73	9	HS238523	HS238523 Homo sapi
c 572	21.4	0.4	79	6	AR125976	AR125976 Sequence	c 645	21.2	0.4	74	6	A35497	A35497 Synthetic 1
c 573	21.4	0.4	79	6	I47296	I47296 Sequence 22	c 646	21.2	0.4	75	4	AF330199	AF330199 Sus scrof
c 574	21.4	0.4	80	9	AF205192	AF205192 Homo sapi	c 647	21.2	0.4	79	1	AX033166	AX033166 Sequence
c 575	21.4	0.4	81	9	AF020775	AF020775 Homo sapi	c 648	21.2	0.4	79	6	AF213170	AF213170 Metagelis
c 576	21.4	0.4	82	9	HUM6PDS03	M26750 Human gluco	c 649	21.2	0.4	79	6	AX033158	AX033158 Sequence

c 650	21.2	0.4	81	14	AF463312	Hepatitis	723	21	0.4	78	6	AX127961	AX127961 Sequence
c 651	21.2	0.4	82	6	AR074463	Sequence	724	21	0.4	78	10	AF265826	AF265826 Mus muscu
c 652	21.2	0.4	82	6	AR081143	Sequence	725	21	0.4	79	6	AX099452	AX099452 Sequence
c 653	21.2	0.4	82	6	AR085340	Sequence	726	21	0.4	79	6	AX233523	AX233523 Sequence
c 654	21.2	0.4	82	6	AR080808	Sequence	727	21	0.4	79	6	AX381268	AX381268 Sequence
c 655	21.2	0.4	82	6	AR104247	Sequence	728	21	0.4	79	10	RNENDARR2	RNENDARR2 Sequence
c 656	21.2	0.4	82	6	AR143511	Sequence	729	21	0.4	80	6	I34337	I34337 R. norvegicu
c 657	21.2	0.4	82	6	AR171415	Sequence	730	21	0.4	80	6	AB010657	AB010657 Homo sapi
c 658	21.2	0.4	82	6	AR171586	Sequence	731	21	0.4	81	6	I19250	I19250 Sequence 22
c 659	21.2	0.4	82	6	AF46118852	Sequence	732	21	0.4	81	6	I19273	I19273 Sequence 45
c 660	21.2	0.4	84	9	HSCR2AA	Homo sapien	733	21	0.4	81	14	AB045015	AB045015 Hepatitis
c 661	21.2	0.4	86	6	AR042809	Sequence	734	21	0.4	81	14	AF390588	AF390588 Hepatitis
c 662	21.2	0.4	86	8	AY06555852	Sequence	735	21	0.4	83	8	SIARRN03	SIARRN03 Stellaria m
c 663	21.2	0.4	87	6	AR007041	Sequence	736	21	0.4	84	6	AR052315	AR052315 Sequence
c 664	21.2	0.4	87	6	AR111065	Sequence	737	21	0.4	84	9	AB010648	AB010648 Homo sapi
c 665	21.2	0.4	87	6	AR208319	Sequence	738	21	0.4	84	10	MUSICAPCH	MUSICAPCH Mus muscu
c 666	21.2	0.4	87	6	I74981	Sequence 18	739	21	0.4	86	3	DDIACMC	DDIACMC P. discoiden
c 667	21.2	0.4	87	9	AF129751	Homo sapi	740	21	0.4	86	6	AR006930	AR006930 Sequence
c 668	21.2	0.4	87	9	SS5093	Ig C gamma	741	21	0.4	86	6	AR110954	AR110954 Sequence
c 669	21.2	0.4	88	9	U00823	Human Beta-	742	21	0.4	86	6	AX114858	AX114858 Sequence
c 670	21.2	0.4	88	17	HSMC3E05	Sequence	743	21	0.4	86	6	I74870	I74870 ompa signal
c 671	21.2	0.4	90	6	AX404025	Sequence	744	21	0.4	87	6	A20577	A20577 ompa signal
c 672	21.2	0.4	90	6	AX429828	Sequence	745	21	0.4	87	6	AR052316	AR052316 Sequence
c 673	21.2	0.4	90	9	D86115	Homo sapien	746	21	0.4	88	6	AR166010	AR166010 Sequence
c 674	21.2	0.4	90	10	MMVAF832J	Mus musculus	747	21	0.4	88	6	AR209981	AR209981 Sequence
c 675	21.2	0.4	92	9	AY006095	Homo sapi	748	21	0.4	90	6	AR152236	AR152236 Sequence
c 676	21.2	0.4	92	9	HSSCN1B56	Sequence	749	21	0.4	90	6	AX000947	AX000947 Sequence
c 677	21.2	0.4	92	14	MCA4184P	Sequence	750	21	0.4	90	6	AX031090	AX031090 Sequence
c 678	21.2	0.4	93	6	AR026962	Sequence	751	21	0.4	90	12	SYNVP7BG2	SYNVP7BG2 Synthetic v
c 679	21.2	0.4	96	6	E07898	DNA fragmen	752	21	0.4	90	14	HSILPF	HSILPF Herpes simp
c 680	21.2	0.4	96	10	AF041916	Mus muscu	753	21	0.4	90	14	HSILPH	HSILPH Herpes simp
c 681	21.2	0.4	98	6	AX287659	Sequence	754	21	0.4	90	14	HSILPH	HSILPH Herpes simp
c 682	21.2	0.4	98	10	MUSFC5F08	Sequence	755	21	0.4	90	14	OV030345	OV030345 Orf virus D
c 683	21.2	0.4	99	3	DPZEN	Mus sapiens	756	21	0.				

796	20.8	0.4	51	6	AX162154	Sequence	869	20.8	0.4	98	6	AX287658	Sequence	AX287658	Sequence
797	20.8	0.4	51	6	AX164995	Sequence	870	20.8	0.4	98	6	AX287664	Sequence	AX287664	Sequence
C 798	20.8	0.4	51	6	AX190100	Sequence	871	20.8	0.4	98	6	AX287665	Sequence	AX287665	Sequence
800	20.8	0.4	54	6	A68171	Sequence	C 872	20.8	0.4	98	6	E36069	Sequence	E36069	Higher-order
801	20.8	0.4	54	6	E07203	Oligonucleo	C 873	20.8	0.4	98	6	I49625	Sequence	I49625	Sequence 42
802	20.8	0.4	54	6	E13400	Sequence	C 874	20.8	0.4	99	5	GGCOL15	Sequence	GGCOL15	Sequence
803	20.8	0.4	54	6	E125365	Sequence	C 875	20.8	0.4	99	5	AR078745	Sequence	AR078745	Sequence
804	20.8	0.4	56	6	AX139459	Sequence	C 876	20.8	0.4	99	6	AR078745	Sequence	AR078745	Sequence
805	20.8	0.4	56	6	AR117941	Sequence	C 877	20.8	0.4	100	6	AR037626	Sequence	AR037626	Sequence
806	20.8	0.4	61	9	AB032803	Sequence	C 878	20.8	0.4	100	6	AR109037	Sequence	AR109037	Sequence
807	20.8	0.4	65	6	A45267	Sequence	C 879	20.8	0.4	100	6	AR166596	Sequence	AR166596	Sequence
808	20.8	0.4	67	6	AR072420	Sequence	C 880	20.8	0.4	100	6	AR178084	Sequence	AR178084	Sequence
C 809	20.8	0.4	67	6	126531	Sequence	C 881	20.8	0.4	100	10	RATMAT01	Sequence	RATMAT01	Sequence
C 810	20.8	0.4	68	6	AR012359	Sequence	C 882	20.6	0.4	21	6	AX095834	Sequence	AX095834	Sequence
C 811	20.8	0.4	69	6	AF218767	Sequence	C 883	20.6	0.4	21	6	AX095835	Sequence	AX095835	Sequence
C 812	20.8	0.4	70	9	HSAL1SN	Sequence	C 884	20.6	0.4	31	6	E11402	Sequence	E11402	Primer 9/1
C 813	20.8	0.4	72	6	A90898	Sequence	C 885	20.6	0.4	37	6	AX097685	Sequence	AX097685	Sequence
814	20.8	0.4	72	6	AR072426	Sequence	C 886	20.6	0.4	37	6	AX166887	Sequence	AX166887	Sequence
C 815	20.8	0.4	72	6	126537	Sequence	C 887	20.6	0.4	37	6	AX167014	Sequence	AX167014	Sequence
C 816	20.8	0.4	73	6	MUSC1A13	Sequence	C 888	20.6	0.4	44	6	E11301	Sequence	E11301	PCR primer
C 817	20.8	0.4	73	6	A08915	Sequence	C 889	20.6	0.4	46	3	LE1KPERB	Sequence	LE1KPERB	Sequence
C 818	20.8	0.4	74	6	HUMGCB2	Sequence	C 890	20.6	0.4	47	6	AX195039	Sequence	AX195039	Sequence
C 819	20.8	0.4	74	6	E02814	Sequence	C 891	20.6	0.4	51	6	AX156814	Sequence	AX156814	Sequence
C 820	20.8	0.4	74	6	E03955	Sequence	C 892	20.6	0.4	51	6	AX157507	Sequence	AX157507	Sequence
C 821	20.8	0.4	75	6	AX112906	Sequence	C 893	20.6	0.4	51	6	AX165548	Sequence	AX165548	Sequence
C 822	20.8	0.4	75	6	AX474734	Sequence	C 894	20.6	0.4	51	6	AX190148	Sequence	AX190148	Sequence
C 823	20.8	0.4	76	6	AX381220	Sequence	C 895	20.6	0.4	51	6	AX203990	Sequence	AX203990	Sequence
C 824	20.8	0.4	76	6	AF040227	Sequence	C 896	20.6	0.4	51	6	AX204000	Sequence	AX204000	Sequence
C 825	20.8	0.4	78	5	AF035411	Sequence	C 897	20.6	0.4	51	6	184400	Sequence	184400	Sequence 1
826	20.8	0.4	78	5	AF035411	Sequence	C 898	20.6	0.4	54	10	AF224107	Sequence	AF224107	Mus muscu
827	20.8	0.4	79	6	A61619	Sequence	C 899	20.6	0.4	54	10	AF224118	Sequence	AF224118	Mus muscu
C 828	20.8	0.4	80	6	AF068162	Sequence	C 900	20.6	0.4	56	6	A38976	Sequence	A38976	Sequence
C 829	20.8	0.4	80	6	117605	Sequence	C 901	20.6	0.4	56	6	AR054372	Sequence	AR054372	Sequence
C 830	20.8	0.4	80	6	AF062767	Sequence	C 902	20.6	0.4	56	6	AX139457	Sequence	AX139457	Sequence
C 831	20.8	0.4	81	3	LIMSREP1	Sequence	C 903	20.6	0.4	56	9	HSYCRBCH	Sequence	HSYCRBCH	Sequence
C 832	20.8	0.4	81	10	AF265744	Mus muscu	C 904	20.6	0.4	57	6	AX167167	Sequence	AX167167	Sequence
C 833	20.8	0.4	81	14	AF221237	Hepatitis	C 905	20.6	0.4	59	6	AR146500	Sequence	AR146500	Sequence
C 834	20.8	0.4	82	6	117606	Sequence	C 906	20.6	0.4	59	6	S78804	Sequence	S78804	Sequence
C 835	20.8	0.4	83	6	A64750	Sequence	C 907	20.6	0.4	60	6	AX404653	Sequence	AX404653	Sequence
C 836	20.8	0.4	83	6	AR161647	Sequence	C 908	20.6	0.4	60	9	AF011625	Sequence	AF011625	Homo sapi
C 837	20.8	0.4	83	11	G36392	Sequence	C 909	20.6	0.4	61	5	AR118194	Sequence	AR118194	Sequence
C 838	20.8	0.4	84	3	LIMSREP1	Sequence	C 910	20.6	0.4	62	5	E26097	Sequence	E26097	Novel DNA f
C 839	20.8	0.4	84	9	HUMD4H08M3	Sequence	C 911	20.6	0.4	62	5	105585	Sequence	105585	Sequence
840	20.8	0.4	84	9	HSU5108	Sequence	C 912	20.6	0.4	64	3	DMW3AS	Sequence	DMW3AS	Sequence
841	20.8	0.4	87	17	HSWC04A06	Sequence	C 913	20.6	0.4	64	5	AX080203	Sequence	AX080203	Sequence
842	20.8	0.4	89	6	AX381436	Sequence	C 914	20.6	0.4	65	4	AF137375	Sequence	AF137375	Sus scrofa
843	20.8	0.4	91	10	AB010320	Sequence	C 915	20.6	0.4	65	6	AX482957	Sequence	AX482957	Sequence
C 844	20.8	0.4	92	6	A44963	Sequence	C 916	20.6	0.4	65	6	AX484851	Sequence	AX484851	Sequence
C 845	20.8	0.4	92	6	AX093633	Sequence	C 917	20.6	0.4	65	6	AR027783	Sequence	AR027783	Sequence
C 846	20.8	0.4	92	6	164454	Sequence	C 918	20.6	0.4	65	8	AR146499	Sequence	AR146499	Sequence
C 847	20.8	0.4	93	6	AR074458	Sequence	C 919	20.6	0.4	66	6	AF265963	Sequence	AF265963	Sequence
C 848	20.8	0.4	93	6	AR081138	Sequence	C 920	20.6	0.4	66	10	MMU84673	Sequence	MMU84673	Sequence
C 849	20.8	0.4	93	6	AR085335	Sequence	C 921	20.6	0.4	66	10	AF265963	Sequence	AF265963	Sequence
C 850	20.8	0.4	93	6	AR088083	Sequence	C 922	20.6	0.4	69	6	AL17956	Sequence	AL17956	Sigmal pept
C 851	20.8	0.4	93	6	AR104242	Sequence	C 923	20.6	0.4	69	6	AX119978	Sequence	AX119978	Sequence
C 852	20.8	0.4	93	6	AR143506	Sequence	C 924	20.6	0.4	69	6	AX138027	Sequence	AX138027	Sequence
C 853	20.8	0.4	93	6	AR171410	Sequence	C 925	20.6	0.4	69	12	AF411478	Sequence	AF411478	Synthetic
C 854	20.8	0.4	93	6	AR171581	Sequence	C 926	20.6	0.4	70	5	LPN28D8X	Sequence	LPN28D8X	Sequence
C 855	20.8	0.4	93	9	HSC1INR2	Sequence	C 927	20.6	0.4	70	6	AR123757	Sequence	AR123757	Sequence
C 856	20.8	0.4	93	10	RATWCR1B	Sequence	C 928	20.6	0.4	71	6	AA2300	Sequence	AA2300	Sequence
C 857	20.8	0.4	94	5	MTACAL18B	Sequence	C 929	20.6	0.4	71	6	AA2301	Sequence	AA2301	Sequence
C 858	20.8	0.4	94	5	MTSAL18SB	Sequence	C 930	20.6	0.4	72	6	AB2289	Sequence	AB2289	Sequence
C 859	20.8	0.4	94	5	MTPEL18SB	Sequence	C 931	20.6	0.4	72	6	AR063726	Sequence	AR063726	Sequence
C 860	20.8	0.4	94	5	MTPEL18SB	Sequence	C 932	20.6	0.4	72	6	AR122149	Sequence	AR122149	Sequence
C 861	20.8	0.4	94	5	AX187852	Sequence	C 933	20.6	0.4	72	6	AR142607	Sequence	AR142607	Sequence
C 862	20.8	0.4	94	6	AX187852	Sequence	C 934	20.6	0.4	72	6	AX150235	Sequence	AX150235	Sequence
C 863	20.8	0.4	94	11	HSPE75H11	Sequence	C 935	20.6	0.4	72	6	AX474798	Sequence	AX474798	Sequence
C 864	20.8	0.4	95	1	CUE133575	Sequence	C 936	20.6	0.4	72	6	130089	Sequence	130089	Sequence
C 865	20.8	0.4	95	6	AR033948	Sequence	C 937	20.6	0.4	72	9	AF189555	Sequence	AF189555	Homo sapi
C 866	20.8	0.4	95	6	AR175081	Sequence	C 938	20.6	0.4	73	9	AF132018	Sequence	AF132018	Homo sapi
C 867	20.8	0.4	95	6	AX032525	Sequence	C 939	20.6	0.4	73	9	AF132490	Sequence	AF132490	Homo sapi
C 868	20.8	0.4	97	6	HUMGCB2	Sequence	C 940	20.6	0.4	73	9	AF132491	Sequence	AF132491	Homo sapi
C 868	20.8	0.4	98	6	AR087170	Sequence	C 941	20.6	0.4	73	9	HSB07HLAI	Sequence	HSB07HLAI	Homo sapi

C 942	20.6	0.4	73	9	HSNLAGM1	AF110259 Homo sapi
C 943	20.6	0.4	73	9	SHSLABKM1	AF110262 Homo sapi
C 944	20.6	0.4	74	6	AR202759	AR202759 Sequence
C 945	20.6	0.4	74	9	HUMNTCOF06	U84645 Human putat
C 946	20.6	0.4	75	6	AX080204	AX080204 Sequence
C 947	20.6	0.4	75	11	G67162	G67162 sv639 Human
C 948	20.6	0.4	77	6	AR063284	AR063284 Sequence
C 949	20.6	0.4	77	6	AR156845	AR156845 Sequence
C 950	20.6	0.4	77	6	AR179022	AR179022 Sequence
C 951	20.6	0.4	77	6	AX147663	AX147663 Sequence
C 952	20.6	0.4	77	6	AX166288	AX166298 Sequence
C 953	20.6	0.4	77	6	AX381197	AX381197 Sequence
C 954	20.6	0.4	77	9	HUMGP60H18	M57563 Human gluc
C 955	20.6	0.4	78	6	AX192757	AX192757 Sequence
C 956	20.6	0.4	79	10	MMTG39	X00357 Mouse germ
C 957	20.6	0.4	80	6	AX480698	AX480698 Sequence
C 958	20.6	0.4	80	6	I00999	I00999 Sequence 2
C 959	20.6	0.4	80	6	I02190	I02190 Sequence 2
C 960	20.6	0.4	81	3	DROHOMEOAC	L08620 D.melanog
C 961	20.6	0.4	81	6	AR031497	AR031497 Sequence
C 962	20.6	0.4	81	6	AR044269	AR044269 Sequence
C 963	20.6	0.4	81	6	AR044270	AR044270 Sequence
C 964	20.6	0.4	81	6	AR130874	AR130874 Sequence
C 965	20.6	0.4	81	6	I35909	I35909 Sequence 14
C 966	20.6	0.4	81	6	I35910	I35910 Sequence 14
C 967	20.6	0.4	81	6	I68882	I68882 Sequence 14
C 968	20.6	0.4	81	14	I68883	I68883 Sequence 15
C 969	20.6	0.4	81	14	AF166738	AF166738 Hepatitis
C 970	20.6	0.4	81	14	AF166745	AF166745 Hepatitis
C 971	20.6	0.4	81	14	AF463306	AF463306 Hepatitis
C 972	20.6	0.4	81	14	AF463307	AF463307 Hepatitis
C 973	20.6	0.4	81	14	AF463313	AF463313 Hepatitis
C 974	20.6	0.4	81	14	AF463314	AF463314 Hepatitis
C 975	20.6	0.4	81	14	AF463318	AF463318 Hepatitis
C 976	20.6	0.4	81	14	AF463319	AF463319 Hepatitis
C 977	20.6	0.4	81	14	AF463320	AF463320 Hepatitis
C 978	20.6	0.4	81	14	AF463321	AF463321 Hepatitis
C 979	20.6	0.4	81	14	AF463322	AF463322 Hepatitis
C 980	20.6	0.4	81	14	AF463323	AF463323 Hepatitis
C 981	20.6	0.4	81	14	AF463324	AF463324 Hepatitis
C 982	20.6	0.4	81	14	AF463326	AF463326 Hepatitis
C 983	20.6	0.4	81	14	AF463327	AF463327 Hepatitis
C 984	20.6	0.4	81	14	AF463328	AF463328 Hepatitis
C 985	20.6	0.4	81	14	AF463331	AF463331 Hepatitis
C 986	20.6	0.4	82	5	AF044810	AF044810 Lampetra
C 987	20.6	0.4	84	6	I11717	I11717 Sequence 3
C 988	20.6	0.4	84	6	I11733	I11733 Sequence 3
C 989	20.6	0.4	84	6	I43849	I43849 Sequence 3
C 990	20.6	0.4	85	4	AF132603	AF132603 Lama glam
C 991	20.6	0.4	85	6	AX136915	AX136915 Sequence
C 992	20.6	0.4	85	6	BD006812	BD006812 Method fo
C 993	20.6	0.4	86	6	A50568	A50568 Sequence 10
C 994	20.6	0.4	86	6	AR147245	AR147245 Sequence
C 995	20.6	0.4	86	6	AX203294	AX203294 Sequence
C 996	20.6	0.4	86	6	AX240998	AX240998 Sequence
C 997	20.6	0.4	87	10	MMH2DK06	X53970 Mouse recom
C 998	20.6	0.4	88	6	AR147061	AR147061 Sequence
C 999	20.6	0.4	88	6	AR147068	AR147068 Sequence
C1000	20.6	0.4	88	6	AR147073	AR147073 Sequence

## ALIGNMENTS

RESULT 1			
LOCUS	AX249344	31 bp	
DEFINITION	Sequence 1423 from Patent WO0166800		
ACCESSION	AX249344		
VERSION	AX249344.1		
KEYWORDS	GI:15863967		
SOURCE	human.		
ORGANISM	Homo sapiens		

31 bp  
T

Linear PAT 28-SEP-2001

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 31)	Cargill, M., Letant, J. S. and Lader, E. S.	Human single nucleotide polymorphisms	Patent: WO 0166800-A 1423 13-SEP-2001.
			WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
		location/qualifiers	
		1	31

BASE COUNT	5 a	15 c	9 g	1 t	1 others
ORIGIN					

Query Match	0.68;	Score 30.6;	DB 6;	Length 31;
Best Local Similarity	96.88;	Pred. No. 2e+05;		
Matches 30; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY 2865 AGGAGGGCGCGCCGACGCCCCCTGCCCAAC 2895
      |||||:|||||
Db 1 AGGAGGGCGCGCCGACGCCCCCTGCCCAAC 31

```

RESULT 2	AX249345	31 bp	DNA	linear	PAT 28-SEP-2001
LOCUS	AX249345				
DEFINITION	Sequence 1424 from Patent WO0166800.				
ACCESSION	AX249345				
VERSION	AX249345.1	GI:15863968			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE 1 (bases 1 to 31)  
AUTHORS Cargill, M., Ireland, J. S. and Lander, E. S.  
TITLE Human single nucleotide polymorphisms  
JOURNAL Patent: WO 0166800-A 124 13-SEP-2001  
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)  
location/Qualifiers  
1. .31

BASE COUNT	5 a	10 c	13 g	2 t	1 others
ORIGIN					

Query Match	0.6%	Score 30.6;	DB 6;	Length 31;
Best Local Similarity	96.8%;	Pred. No. 2e+05;		
Matches 30; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0

Qy	3032	AGCCGCAGATGCGCTGGGCCACGGCGTGAG	306
		:	
Db	1	AGCCGCAGATGCGCTGGGCCACGGCGTGAG	31

LOCUS	DEFINITION	SEQUENCE	ACCESSION	VERSION	GI	78 bp	DNA	1 linear	PAT 17-OCT-2001
AR159551	AR159551	Sequence	11	from patent	US 6251589.				
AR159551	AR159551								
AR159551.1	AR159551.1				GI:16222238				

ORGANISM	Unknown.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Unclassified.  
REFERENCE 1 (bases 1 to 78)

**AUTHORS**  
Tsuji, S. and Sanpei, K.

TITLE Method for diagnosing spinocerebellar ataxia type 2 and primers therefor

JOURNAL Patent: US 6251589-A 11 26-JUN-2001

source	1. .78
/organism="unknown"	



Db	7	CAGCCCGGCACTTGGCTTCACACTTGCGCAGATCAGGCCAGGCCCCCTCCGTGCCTCCC	66
QY	3091	TCCGAGGCGCTGGCCCTGCTGCTGTGCCCGC	3123
Db	67	GTCGCCGACCTGCTCCACAGCCGGGCGTGGCGGC	99
RESULT 10			
LOCUS	173148		99 bp DNA
DEFINITION	Sequence 10 from patent US 5686239.		Linear PAT 03-APR-1998
ACCESSION	173148		
VERSION	173148.1 GI:3009287		
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 99)		
AUTHORS	Reyes,G.R., Tam,A.W. and Yarbough,P.O.		
TITLE	Hepatitis E virus peptides and methods		
JOURNAL	Patent: US 5686239-A 10 11-NOV-1997;		
FEATURES	Location/Qualifiers		
source	1..99		
BASE COUNT	12 a 46 c 29 g	12 t	
ORIGIN			
Query Match	0.5%, Score 27.4; DB 6; Length 99;		
Best Local Similarity	55.9%; Pred. No. 7.8e+05;		
Matches	52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;		
QY	3031 CAGCCGACGATGCGCTGGCCACGCGCTGAAGAAGCGCCAGCGACCCTGGCGGACAGC	3090	
Db	7 CAGCCCGGCACTTGGCTTCACACTTGCGCAGATCAGGCCAGGCCCCCTCCGTGCCTCCC	66	
QY	3091 TCCGAGGCGCTGGCCCTGCTGCTGTGCCCGC	3123	
Db	67 GTCGCCGACCTGCTCCACAGCCGGGCGTGGCGGC	99	
RESULT 11			
LOCUS	PTU18443/c	98 bp DNA	PRI 13-NOV-1995
DEFINITION	Pan troglodytes pre-Alu-integration site clone Chimpa62.		
ACCESSION	U18443		
VERSION	U18443.1 GI:841185		
KEYWORDS			
SOURCE	Pan troglodytes.		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.		
TITLE	1 (bases 1 to 98)		
JOURNAL	Arcole,S.S., Deininger,P.L. and Batzer,M.A.		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 98)		
TITLE	Batzer,M.A., Kilroy,G.E., Richard,P.E., Shaikh,T.H., Desselte,T.D.,		
JOURNAL	Hoppens,C.L. and Deininger,P.L.		
REFERENCE	Structure and variability of recently inserted Alu family members		
AUTHORS	Nucleic Acids Res. 18 (23), 6793-6798 (1990)		
TITLE	91088248		
JOURNAL	2175877		
REFERENCE	3 (bases 1 to 98)		
AUTHORS	Arcole,S.S., Wang,Z., Weber,J.L., Deininger,P.L. and Batzer,M.A.		
TITLE	Alu repeats: a source for the genesis of primate microsatellites		
JOURNAL	Genomics 29 (1), 136-144 (1995)		
MEDLINE	96079101		
PUBMED	8530063		
REFERENCE	4 (bases 1 to 98)		
AUTHORS	Batzer,M.A., Rubin,C.M., Hellmann-Bumberg,U., Alegria-Hartman,M.,		
TITLE	Loefflang,E.P., Sten,J.D., Bazan,H.A., Shaikh,T.H., Deininger,P.L.		
JOURNAL	and Schmid,C.W.		
REFERENCE	Dispersion and insertion polymorphism in two small subfamilies of		



PROCEDURES - NOV 1971

Copied from PCI US2003092 on 04-03-2004

Best Local Similarity	66.18;	Pred. No.	9.2e+05;
Matches	39;	Conservative	0;
		Mismatches	20;
		Indels	0;
		Gaps	0;

**Q7**    3473 GCCAGCAGTTCATGCCCCCTCGAGCAGACGCCCTGGCCCCGAGGGCAGCAAAACCACACTGCCC    3531  
         | ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
**Db**     29 GGCAGGAGTGCCATTCCACGAGTAACCGGACGGGAGCGCAGGAAGTCGGAGGTGCC    87

RESULT 15	AX031011/c	LOCUS	AX031011	99 bp	DNA	linear	PAT 20-SEP-2000
DEFINITION	Sequence 13 from Patent WO9806830.						
ACCSSION	ax031011						

BASE COUNT	20 a	38 c	22 g	19 t
ORIGIN				

Query Match	0.5%;	Score 27;	DB 6;	Length 99;
Best Local Similarity	66.1%;	Pred. No. 9.2e+05;		
Matches 39; Conservative	0;	Mismatches 20;	Indels 0;	Gaps 0;

**QY** 3473 GCGACAGTTCATGCCCCCTCGACGACGCCCTGCCCCGAGGGCAGCAAAACCACCTGCC 3531  
| | | | | | | | | | | | | | | | | | | | | |  
**Db** 75 GGCAGAGATCCCATTCACGAGTAACGCCGAGCGGGAGCGCAGGAATCGAGGTGCC 17

RESULT 16			
AX031041/c			
LOCUS	AX031041	99 bp	DNA
DEFINITION	Sequence 13 from Patent WO9806830.		linear
ACCESSION	AX031041		
VERSION	AX031041.1	GI:10278434	

BASE COUNT	20 a	38 c	22 g	19 t
ORIGIN	/organism="unidentified" /db_xref="taxon:32644"			

Query Match	0.5%	Score 27;	DB 13;	Length 99;
Best Local Similarity	66.1%	Pred. No. 9.2e+05;		
Matches 39; Conservative	0;	Mismatches 20;	Indels 0;	Gaps 0;

Oy 3473 GCCAGCAGTTCATGCGCCCTGCAGACAGCCCTGCCCGAGGGCAGCAAAACCGACTGCC 3531D  
+ ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 75 GCGCAGAGTGCCATTCCACAGAGTAACGCGGAGCGCACAGAATCGGAGGTGCC 17

RESULT 17			
HSPE06C12			
LOCUS	82 bp	DNA	linear
HSPE06C12			STS 21-MAY-1998

**DEFINITION** H.sapiens flow-sorted chromosome 1 HindIII fragment, SC1pE06C12 sequence tagged site.

ACCESSION	AL0009338
VERSION	AL0009338.1
KEYWORDS	STS, single read.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 82)	Gregory, S., Kettleborough, R., Langford, C., Ross, M.T. and Hunt, S.E.	Direct Submission	Submitted (03-DEC-1997)	E-mail contact: humquerry@sanger.ac.uk Vector: pBS1818+ Marker sts828963 (Primer A: CAGCGGAGGCTCATGTC; Primer B: GCTTCGGCAGCACTCATTTAC; amplicon size : 82 bp) was mapped to chromosome 1 using Radiation Hybrid panel Genebridge 4 (GB4).

FEATURES	
SOURCE	location/qualities
	1. .82
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/clone="SC1PE06C12"
	/sex="Female"
	/tissue_type="EBV lymphoblastoid cell line"
	/clone_lib="SC1PE"
	/dev_stage="adult"
BASE COUNT	25 a 20 c 20 g 17 t
ORIGIN	

Query Match	0.58;	Score 26.6;	DB 11;	Length 82;
Best Local Similarity	71.48;	Pred. No. 1.1e+06;		
Matches 35; Conservative	0;	Mismatches 14;	Indels 0;	Gaps 0

```

QY 1742 TGAACACACTTGAGATCTCACACTGGAGAGAAACATACGCTGTGA 1790
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18 TCATCACACACCGAGAAATTCACACTGGAGAGAAAGCTTATGAGTGTAA 66

```

RESULT 18	
SSGGRTP1	
LOCUS	SSGGRTP1 88 bp DNA linear MAM 16-DEC-1994
DEFINITION	X. scrofa DNA for gamma-glutamyl transpeptidase (5' end).
ACCESSION	X16533
VERSION	X16533.1 GI:1954
KEYWORDS	gamma-glutamyl transpeptidase.
SOURCE	Xus scrofa.

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 88)	Papandrikopoulou, A., Frey, A. and Gassen, H.G.	Cloning and expression of gamma-glutamyl transpeptidase from isolated porcine brain capillaries	Eur. J. Biochem. 183 (3), 693-698 (1989)

PUBMED	2476308
COMMENT	See X55020 for continued cds (mRNA sequence).
FEATURES	Location/Qualifiers
source	1..88

```

CDS
    /db_xref="taxon:9623-
    36.">88
    /EC_number="2.3.2.2"
    /codon_start=1
    /product="gamma-glutamyltranspeptidase"
    /protein_id="CAA34356.1"
    /db_xref="gi:602909"
    /db_xref="SWISS-PROT:P20735"
    /translation="WKKRYLLLAALAAVALLL"
BASE COUNT
    13 a      32 c      24 g      19 t
ORIGIN

```

Query Match 0.5%; Score 26.6; DB 4; Length 88;  
 Best Local Similarity 60.3%; Pred. No. 1.1e+06;  
 Matches 44; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 2878 CCGACGCCCTCCGACCAATGAGAGATGACCTGAAGACGCCGCTGCTGCTGG 2937  
 Db 12 CCTCTCTCCCTCCGACCAATGAGAGATGACCTGAAGACGCCGCTGCTGCTGG 71

QY 2938 GATGCCCTCGAGC 2950  
 Db 72 GTGGCCCTAGTGC 84

RESULT 19  
 ARI59550  
 LOCUS ARI59550 75 bp DNA linear PAT 17-OCT-2001  
 DEFINITION Sequence 10 from patent US 6251589.  
 ACCESSION ARI59550  
 VERSION ARI59550.1 GI:16222236  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

REFERENCE  
 1 (bases 1 to 75)  
 Tsuji,S. and Sanpei,K.  
 TITLE Method for diagnosing spinocerebellar ataxia type 2 and primers therefor  
 JOURNAL Patent: US 6251589-A 10 26-JUN-2001;  
 FEATURES  
 Location/Qualifiers  
 source 1..75  
 /organism="unknown"

BASE COUNT 23 a 29 c 23 g 0 t

Query Match 0.5%; Score 26.4; DB 6; Length 75;  
 Best Local Similarity 65.0%; Pred. No. 1.2e+06;  
 Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 2586 CAACAGAAAGGACAGCAGCGCCACACATCAGCTCGCTACTGAGCAGCGCCGCTC 2645  
 Db 13 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 72

RESULT 20  
 H1AS1A  
 LOCUS H1AS1A 100 bp DNA linear VRL 02-AUG-1993  
 DEFINITION Human herpes simplex virus type 1 (HSV-1) 'a' sequence, clone G-2.  
 ACCESSION M13884  
 VERSION M13884.1 GI:330052  
 KEYWORDS  
 SOURCE Human herpes simplex virus type 1 (strain F) DNA, passed in rabbit skin cells, clone G-2.  
 ORGANISM human herpesvirus 1  
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.

REFERENCE  
 1 (bases 1 to 100)  
 Deiss,L.P., Chou,J. and Frenkel,N.  
 TITLE Functional domains within the a sequence involved in the cleavage-packaging of herpes simplex virus DNA  
 JOURNAL J. Virol. 59 (3), 605-618 (1986)  
 MEDLINE 86281869  
 PUBMED 3016323  
 FEATURES  
 Location/Qualifiers  
 source 1..100  
 /organism="human herpesvirus 1"  
 /db\_xref="taxon:10298"

BASE COUNT 4 a 55 c 34 g 7 t

Query Match 0.5%; Score 26.4; DB 14; Length 100;  
 Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
 Matches 48; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 2742 CCCCATCTCCACGACGACCTCGCGCCCTCCAGCGAAGCCAGACGAGCGCTGCC 2801  
 Db 1 CCGCGGCGCCCGCCCGCCAGCGCCCGCGCGCGCGAGCGCGCCCGGAGCGCGCGCGC 60

QY 2802 CAGCCTGCTCAGGCTCAGCGCCGCG 2825  
 Db 61 TTTTTCGCGCGCGCGCCCGCGCGC 84

RESULT 21  
 AF189453  
 LOCUS AF189453 63 bp mRNA linear PRI 01-FEB-2000  
 DEFINITION Homo sapiens clone C45U T-cell receptor beta chain (TCRBV67) mRNA, partial cds.  
 ACCESSION AF189453  
 VERSION AF189453.1 GI:6841773  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 63)  
 Soudeyns,H., Champagne,P., Holloway,C.L., Silvestri,G.U., Ringuette,N., Samson,J., Lapointe,N. and Sekaly,R.P.  
 TITLE Transient T cell receptor beta-chain variable region-specific expansions of CD4+ and CD8+ T cells during the early phase of pediatric human immunodeficiency virus infection: characterization of expanded cell populations by T cell receptor phenotyping  
 J. Infect. Dis. 181 (1), 107-120 (2000)

JOURNAL MEDLINE 20087004  
 PUBMED 10608757

REFERENCE  
 2 (bases 1 to 63)  
 Soudeyns,H., Champagne,P., Holloway,C.L., Silvestri,G.U., Ringuette,N., Samson,J., Lapointe,N. and Sekaly,R.P.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-SEP-1999) Laboratoire d'immunologie, IRCM, 110 avenue Des Pins ouest, Montreal, Quebec H2W 1R7, Canada

FEATURES  
 source 1..63  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="C45U"  
 <1..>63  
 /gene="TCRBV67"  
 <1..>63  
 /gene="TCRBV67"  
 /note="V(D)J region"  
 /codon\_start=1  
 /product="T-cell receptor beta chain"  
 /protein\_id="AAFP29240.1"  
 /db\_xref="GI:6841774"  
 /translation="QQEDSAVYLCASSIVGYEYF"

BASE COUNT 12 a 17 c 20 g 14 t

Query Match 0.5%; Score 26.2; DB 9; Length 63;  
 Best Local Similarity 63.5%; Pred. No. 1.3e+06;  
 Matches 40; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 3328 CTCGCGAGACGCTGTGCGATTTAAATCCACAGACCAAGCGATGACGAGCAGC 3387  
 Db 1 CAGCAGGAGGACTCGCGCGTGTCTCTGTGCGAGCAGCTTAGTGGGTACAGCAGTAC 60

QY 3388 TTC 3390  
 Db 61 TTC 63

RESULT 22  
 RRG99  
 LOCUS RRG99 99 bp DNA linear ROD 25-APR-1993  
 DEFINITION R.rattus (Sprague Dawley) glucocorticoid receptor gene,

ACCESSION	partial(99bps).	
VERSION	X69669	
KEYWORDS	X69669.1 GI:296700	
SOURCE	glucocorticoid receptor; polyglutamine tract.	
ORGANISM	Rattus rattus	
REFERENCE	Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae;	
AUTHORS	Rattus	
JOURNAL	1 (bases 1 to 99)	
REFERENCE	Gearing, K.L.	
AUTHORS	Direct Submission	
TITLE	Submitted (26-NOV-1992) K.L. Gearing, Dept. Medical Nutrition,	
JOURNAL	Karolinska Institute, Novum, Huddinge University Hospital, S-14186	
MEDLINE	Huddinge, SWEDE	
PUBMED	2 (bases 1 to 99)	
COMMENT	Gearing, K.L., Gustafsson, J.A. and Okret, S.	
FEATURES	Heterogeneity in the polyglutamine tract of the glucocorticoid receptor from different rat strains	
source	Nucleic Acids Res. 21 (8), 2014 (1993)	
Related sequences: M14053 & Y00489.		
Location/Qualifiers		
1. 99		
/organism="Rattus rattus"		
/strain="Sprague Dawley"		
/db_xref="taxon:10117"		
/clone="B"		
/tissue_type="liver"		
/clone_idb="PCR products"		
/dev_stage="adult"		
BASE COUNT	33 a 31 c 26 g 9 t	
ORIGIN		
Query Match	0.5%; Score 26.2; DB 10; Length 99;	
Best Local Similarity	67.3%; Pred. No. 1.3e+06;	
Matches	37; Conservative 0; Mismatches 18; Indels 0; Gaps 0;	
OY	2589 CAGAGGAGCAGCGCGCAGACCATCAGCTCGCCTACCTGAGCAGCGCGGC 2643	
Db	22 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 76	
RESULT 23		
LOCUS	AB032281 99 bp DNA linear VRL 25-JAN-2000	
DEFINITION	TT virus DNA, untranslated region, isolate:At-TTV5-I.	
ACCESSION	AB032281	
VERSION	AB032281.1 GI:6751845	
KEYWORDS		
SOURCE	TT virus (isolate:At-TTV5-I, Specific-host:Notes trivirgatus) DNA.	
ORGANISM	TT virus	
REFERENCE	1 (sites)	
AUTHORS	Viruses; ssDNA viruses; unclassified ssDNA viruses.	
JOURNAL	Okamoto, H., Fukuda, M., Tawara, A., Nishizawa, T., Itoh, Y.,	
TITLE	Hayasaka, I., Tsuda, F., Tanaka, T., Miyakawa, Y. and Mayumi, M.	
KEYWORDS	Species-specific TT viruses and cross-species infection in nonhuman	
ORGANISM	primates	
REFERENCE	J. Virol. 74 (3), 1132-1139 (2000)	
AUTHORS	2 (bases 1 to 99)	
JOURNAL	Okamoto, H.	
TITLE	Direct Submission	
KEYWORDS	Submitted (10-SEP-1999) Hiroaki Okamoto, Jichi Medical School,	
ORGANISM	Immunology Division 3311-1 Yakushiji, Minamikawachi-machi,	
REFERENCE	Kawachi-gun, Tochigi 329-0498, Japan, E-mail:hokamoto@jichi.ac.jp,	
AUTHORS	Tel:81-285-58-7404, Fax:81-285-44-1357)	
JOURNAL	Location/Qualifiers	
TITLE	1. 99	
KEYWORDS	/organism="TT virus"	
ORGANISM	/isolate="At-TTV5-I"	
REFERENCE	/specific-host="Notes trivirgatus"	





MMAPLB2/c  
LOCUS MMAPLB2 88 bp DNA linear ROD 16-JUL-1995  
DEFINITION Mus musculus apolipoprotein B mRNA-editing component 1 (ApoBc1) gene, exon 2.  
ACCESSION U21945  
VERSION U21945.1 GI:899495  
KEYWORDS 2 of 8  
SEGMENT mouse.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS 1 (bases 1 to 88)  
Nakamura, M., Oka, K., Krushkal, J., Kobayashi, K., Yamamoto, M., Li, W. H. and Chan, L.  
TITLE Alternative mRNA splicing and differential promoter utilization determine tissue-specific expression of the apolipoprotein B mRNA-editing protein (ApoBc1) gene in mice. Structure and evolution of ApoBc1 and related nucleoside/nucleotide deaminases  
JOURNAL J. Biol. Chem. 270 (22), 13042-13056 (1995)  
MEDLINE 95286585  
PUBMED 7768898  
REFERENCE 2 (bases 1 to 88)  
AUTHORS Nakamura, M.  
TITLE Direct Submission  
JOURNAL Submitted (28-FEB-1995) Makoto Nakamura, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
FEATURES  
source  
1. 88  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="6"  
<1. 10  
/gene="ApoBc1"  
/number=2  
11. 78  
/gene="ApoBc1"  
/number=2  
/evidence=experimental  
79. >88  
/note="ApoBc1"  
/note="distance between exon 2 and exon 3 is approximately 3.7 kbp"  
/number=1  
BASE COUNT 16 a 22 c 21 g 29 t  
ORIGIN  
Query Match 0.5%; Score 25.2; DB 10; Length 88;  
Best Local Similarity 62.9%; Pred. No. 2e+06;  
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
QY 4018 TACCTGCTCACCACTCTCGGCGACAGCATGACACCCGGGGGCGCGCCGCGT 4077  
DB 81 TACTGTGCTCATGATGCTCGGAGCTAATGTGAAGATCCAGAGACAATCAAT 22  
QY 4078 CA 4079  
DB 21 CA 20  
RESULT 34  
AX320853 93 bp DNA linear PAT 14-DEC-2001  
LOCUS AX320853  
DEFINITION Sequence 23 from Patent WO0183736.  
ACCESSION AX320853  
VERSION AX320853.1 GI:17902404  
KEYWORDS Hepatitis C virus.  
SOURCE Hepatitis C virus.  
ORGANISM Hepatitis C virus.  
REFERENCE Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae; Hepacivirus.  
AUTHORS 1  
Pellerin, C. and Kukolj, G.

TITLE Internal de novo initiation sites of the hcv ns5b polymerase and use thereof  
JOURNAL Patent: WO 0183736-A 23 08-NOV-2001;  
BOEHRINGER INGELHEIM (CANADA) LTD. (CA)  
FEATURES  
source  
1. 93  
/organism="Hepatitis C virus"  
/db\_xref="taxon:11103"  
Location/Qualifiers  
BASE COUNT 2 a 7 c 2 g 82 t  
ORIGIN  
Query Match 0.5%; Score 25.2; DB 6; Length 93;  
Best Local Similarity 55.8%; Pred. No. 2e+06;  
Matches 48; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 4861 TTGTGTTTGGCTGTTTCTTGTCTGATGATTTTACCAATCATCTACCTACTG 4920  
DB 8 TTTTGTGTTTGGCTGTTTCTTGTCTGATGATTTTACCAATCATCTACCTACTG 67  
QY 4921 AGATGTTTCAATTAATTCCTTTT 4946  
DB 68 TTTCCTTTTCTTTTATATTCCTTCT 93  
RESULT 35  
AX008878 94 bp DNA linear PAT 06-SEP-2000  
LOCUS AX008878  
DEFINITION Sequence 3 from Patent WO964585.  
ACCESSION AX008878  
VERSION AX008878.1 GI:9996300  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 94)  
Kleim, J. P., Sibenhorn, G., Heitsch, H. and Riess, G.  
TITLE Novel angiotensin receptor, production and use thereof  
JOURNAL Patent: WO 964585-A 3 16-DEC-1999;  
KLEIM JOERG PETER (DE); SIBENHORN GERTRUD (DE); HEITSCH HOLGER (DE); RIESS GOENTHER (DE); HOECHST MARION ROUSSEL DE GMBH (DE)  
FEATURES  
source  
1. 94  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 30 a 24 c 14 g 26 t  
ORIGIN  
Query Match 0.5%; Score 25.2; DB 6; Length 94;  
Best Local Similarity 57.7%; Pred. No. 2e+06;  
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 138 TGTGAGGAGAAAGCCGTCCTCCAGACACCACTTGAATGAGATGAAGTCTGACA 197  
DB 17 TATTAGCAAAATGAGTAGTACATATTACACTCTAATTTGATGAAGACTCCAGTGC 76  
QY 198 GACTTATCACAGAGAG 215  
DB 77 TACTACACACGAGAG 94  
RESULT 36  
AL841018 96 bp DNA linear STS 17-JUL-2002  
LOCUS AL841018  
DEFINITION Arabidopsis thaliana transposon insertion STS SM\_3.28438, sequence tagged site.  
ACCESSION AL841018  
VERSION AL841018.1 GI:21911912  
KEYWORDS STS; STS, sequence tagged site.  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 Clarke,J.H., Bowles B., Carter,J., Hart,D., McCullagh,B.,
JOURNAL	Murphy,G., Langham,S., Learys,C., Jones,U.D.G. and Bevan,M.
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 96)
TITLE	Clarke,J.H.
JOURNAL	Submitted (15-JUL-2002) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK
COMMENT	AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GR a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3'end of the transposon, _5 denotes a sequence derived from the 5'end of the transposon BBRC GARNet, ARIS Project On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock code: N115926.
FEATURES	location/Qualifiers
SOURCE	1..96
	/organism="Arabidopsis thaliana"
	/variety="Landsberg erecta"
	/db_xref="taxon:3702"
	/clone="AC007138"
STS	1..96
	/standard_name="SM_3.28438"
BASE COUNT	11 a 8 c 56 g 21 t
ORIGIN	
Query Match	0.5%; Score 25.2; DB 11; Length 96;
Best Local Similarity	57.7%; Pred. No.2e+06;
Matches	45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Oy	2728 GCCGACTCTTCAAGACCCTTCACCGAGCCCTGCCGCACGAGCCAGC 2787
Db	87 GCCGACTCTCACCTCTCAACACACACACCGCCCTCCGGTCGAACACTACCCAGC 28
Oy	2788 AGCGACGCCGTGCCAGC 2805
Db	27 TCCTCCACGCGCTCCAC 10
RESULT 37	
LOCUS	AF235063 100 bp DNA linear STS 02-OCT-2000
DEFINITION	Mus musculus chromosome 7 STS, sequence tagged site.
ACCESSION	AF235063
VERSION	AF235063.1 GI:10443758
KEYWORDS	STS.
SOURCE	Mus musculus.
ORGANISM	Mus musculus. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	Wines,M.E., Tiffany,A.M. and Holdener,B.C.
AUTHORS	Physical localization of the mouse aryl hydrocarbon receptor nuclear translocator-2 (Arlt2) gene within the c12k deletion Genomics 51 (2), 223-232 (1998)
JOURNAL	98390185
MEDLINE	9722945
PUBMED	9722945
REFERENCE	2 (bases 1 to 100)
AUTHORS	Wines,M.E.
JOURNAL	Direct Submission Submitted (16-FEB-2000) Biochemistry and Cell Biology, SUNY at Stony Brook, 345 Life Sciences Bldg., Stony Brook, NY 11794-5215,
COMMENT	USA This sequence is from the distal (73) end of a 17.0 kb pBSKII+ NotI subclone from the Sp6 end of Research Genetics BAC 174A19. Sequence is at the 5' end of the Arlt2 gene.
FEATURES	location/Qualifiers
SOURCE	1..100
	/organism="Mus musculus"

STGS	/strain="1298v"
	/db_xref="taxon:10090"
	/chromosome="7"
	<1..>100
BASE COUNT	5 a 51 c 32 g 12 t
ORIGIN	
Query Match	0.5%; Score 25.2; DB 11; Length 100;
Best Local Similarity	54.3%; Pred. No.2e+06;
Matches	51; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
Oy	2934 CGGGGATCCCCTGACGCTGGCGGCGCCCTCCAGTTCAAGCCCCGAGAGTGCCAG 2993
Db	100 CCGGTTCCCTCTTCCGGGAAGGACCGGACACCGGGGGGGGGCGAGGCGC 41
Oy	2994 CCACGGGGAGCCCGACGGCTACGGGGCGCCGCCAC 3027
Db	40 GGAAAGGGGAGCCCGCGCGGGCCAGACCGCCGC 7
RESULT 38	
LOCUS	HSA270365 25 bp DNA linear PRI 26-JUL-2000
DEFINITION	Homo sapiens DNA-binding protein (GLI3) sense primer.
ACCESSION	AJ270365
VERSION	AJ270365.1 GI:9557940
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiade; Homo.
REFERENCE	1 (bases 1 to 25)
AUTHORS	Palm,K., Salin-Nordstrom,T., Levesque,M.F. and Neuman,T.
TITLE	Fetal and adult human CNS stem cells have similar molecular characteristics and developmental potential Brain Res. Mol. Brain Res. 78 (1-2), 192-195 (2000)
JOURNAL	20351569
MEDLINE	10891600
PUBMED	10891600
REFERENCE	2 (bases 1 to 25)
AUTHORS	Palm,K.
JOURNAL	Direct Submission Submitted (04-OCT-1999) Surgery, Cedars Sinai Medical Center, 8700 Beverly Blvd., Los Angeles, CA 90048, US
COMMENT	Related entry: M57609.
FEATURES	location/Qualifiers
SOURCE	1..25
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
misc_feature	1..25
	/note="PCR sense primer for DNA-binding protein (GLI3)"
BASE COUNT	5 a 8 c 6 g 6 t
ORIGIN	
Query Match	0.5%; Score 25; DB 9; Length 25;
Best Local Similarity	100.0%; Pred. No.2.3e+06;
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1031 GCAGCTCTTCAGCAAGTGCCTTA 1055
Db	1 GCAGCTCTTCAGCAAGTGCCTCTTA 25
RESULT 39	
LOCUS	AX039308 98 bp DNA linear PAT 18-NOV-2000
DEFINITION	Sequence 9 from Patent WO0063373.
ACCESSION	AX039308
VERSION	AX039308.1 GI:11229413
KEYWORDS	.
SOURCE	synthetic construct.
ORGANISM	synthetic construct
	artificial sequences.



us-09-910-185-3.max.rge

ad prote

---

aps  
0;

2627

---

26

---

NOV - 2000

---

---

---

0;aps

2627	
26	

---

---



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2002, 00:29:46 : Search time 638 Seconds

(without alignments)  
17843.027 Million cell updates/sec

Title: US-09-910-185-3

Perfect score: 5055  
Sequence: 1 cgatactacgtagggcattt.....accctctttaaanaaaaa 5055

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database: N.Geneseq\_101002.\*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No	Score	Query %	Match Length	DB ID	Description
1	60	1.2	60	24	ABN34264
2	55.4	1.1	65	24	ABN52341
3	31	0.6	31	22	AA130935
4	31	0.6	31	22	AA130936
5	29.6	0.6	78	19	AAV17231
6	28	0.6	69	19	AAV17234
7	28	0.6	78	19	AAV17232
8	28	0.6	78	19	AAV17233
9	28	0.6	88	22	AA65409

10	27.6	0.5	75	22	AA513081	DNA encoding zinc
11	27.6	0.5	90	24	ABK36865	Human DNA encoding
12	27.6	0.5	95	21	AAA97078	G5 synthetic gene
13	27.4	0.5	99	17	AA727118	Hepatitis E virus
14	27.4	0.5	99	18	AA169668	Hepatitis E virus
15	26.8	0.5	51	22	AA128393	Human SNP oligonuc
16	26.8	0.5	51	22	AA128394	Human SNP oligonuc
17	26.4	0.5	75	19	AAV17230	SCA2 gene CAG repe
18	26.2	0.5	51	22	AA126919	Human SNP oligonuc
19	26.2	0.5	89	19	AAV68382	Human SNP oligonuc
20	26.2	0.5	90	20	AAV33913	Clone #8 fragment
21	26.2	0.5	90	20	AAV33913	Human tryptase-I a
22	26.2	0.5	90	21	AAV19478	Human tryptase-I p
23	26.2	0.5	90	21	AAV33356	Low adenosine anti
24	26	0.5	66	18	AAV78908	Poly glutamine rep
25	25.8	0.5	86	22	AA65408	Oligonucleotide B6
26	25.8	0.5	41	22	AAH73705	Human zinc finger
27	25.8	0.5	65	24	ABN52286	Mouse spliced tran
28	25.8	0.5	74	24	AAV24301	der 22 DNA related
29	25.8	0.5	84	21	AAV21832	Human secreted pro
30	25.6	0.5	97	21	AAA36669	RSV and PIV fusion
31	25.6	0.5	60	24	ABN36458	Human spliced tran
32	25.6	0.5	17	17	AAV33692	Nucleic acid encod
33	25.6	0.5	84	17	AAV05454	DNA for immunodu
34	25.6	0.5	92	18	AAV00492	DNA for immunodu
35	25.6	0.5	93	18	AAV05448	DNA for immunodu
36	25.6	0.5	93	18	AAV05448	DNA for immunodu
37	25.6	0.5	93	18	AAV05458	DNA for immunodu
38	25.6	0.5	93	18	AAV05459	DNA for immunodu
39	25.6	0.5	96	18	AAV05456	DNA for immunodu
40	25.6	0.5	96	18	AAV05456	Recombinant adeny
41	25.6	0.5	96	24	ABK43215	Human foetal liver
42	25.6	0.5	98	22	AAV27331	Human brain expres
43	25.6	0.5	98	22	AAV20757	Human brain expres
44	25.6	0.5	98	22	AAV6905	Human bone marrow
45	25.4	0.5	60	24	ABN33803	Probe #21429 used
46	25.4	0.5	63	15	AAQ57443	Human spliced tran
47	25.4	0.5	86	21	AAV05101	D-3-phosphoglyceri
48	25.4	0.5	90	18	AAV76105	Human secreted pro
49	25.2	0.5	99	21	AAV32898	Human tryptase-I a
50	25.2	0.5	50	22	AA128392	Streptococcus pyog
51	25.2	0.5	51	22	AA127313	Human SNP oligonuc
52	25.2	0.5	65	24	ABN27846	Human SNP oligonuc
53	25.2	0.5	83	22	ABK69965	Rat spliced trans
54	25.2	0.5	83	22	ABK69965	Human foetal liver
55	25.2	0.5	83	22	ABK6801	Probe #15267 for g
56	25.2	0.5	83	22	AAK18174	Human brain expres
57	25.2	0.5	83	22	AAK4067	Human bone marrow
58	25.2	0.5	83	22	AA150077	Probe #18763 used
59	25.2	0.5	83	24	ABN18300	Human genome-deri
60	25.2	0.5	86	21	ABN50869	Modified erythro
61	25	0.5	94	21	AAZ43764	Human PCR deriv
62	25	0.5	25	24	ABT03608	Human GII-3 gene p
63	25	0.5	98	22	AA65404	Oligonucleotide B6
64	24.8	0.5	100	22	AA65405	Oligonucleotide B6
65	24.8	0.5	50	22	AA128390	Human SNP oligonuc
66	24.8	0.5	50	22	AA128391	Human SNP oligonuc
67	24.8	0.5	77	22	AA65410	Human SNP oligonuc
68	24.8	0.5	79	22	AA65410	Oligonucleotide B6
69	24.8	0.5	79	22	ABA70065	Human foetal liver
70	24.8	0.5	79	22	ABA36864	Human brain expres
71	24.8	0.5	79	22	AAK18276	Probe #15330 for g
72	24.8	0.5	79	22	AAK44174	Human bone marrow
73	24.8	0.5	79	22	AA150181	Probe #18867 used
74	24.8	0.5	79	22	AA65411	Oligonucleotide B6
75	24.8	0.5	79	22	ABN58409	Oligonucleotide B6
76	24.8	0.5	84	15	AAQ54074	Human genome-deri
77	24.8	0.5	84	15	AAQ54074	DNA encoding a bon
78	24.8	0.5	86	22	AA65408	Oligonucleotide B6
79	24.8	0.5	86	22	AA65409	Oligonucleotide B6
80	24.8	0.5	91	21	AA65409	Human secreted pro
81	24.8	0.5	92	22	AA65406	Oligonucleotide B6
82	24.8	0.5	92	22	AA65407	Oligonucleotide B6
83	24.8	0.5	94	22	AA65407	Oligonucleotide B6
84	24.8	0.5	98	22	AA65407	Oligonucleotide B6
85	24.8	0.5	98	22	AA65407	Oligonucleotide B6
86	24.6	0.5	100	22	AA65405	Oligonucleotide B6
87	24.6	0.5	54	24	ABN5896	Oligonucleotide B6
88	24.6	0.5	54	24	ABN5896	Oligonucleotide B6
89	24.6	0.5	62	21	AA612223	Human secreted pro

C	83	24.6	0.5	92	22	AAC65406	Oligonucleotide B6
C	84	24.6	0.5	94	22	AAC65407	Oligonucleotide B6
C	85	24.6	0.5	97	21	AAA65444	Oligonucleotide SE
C	86	24.6	0.5	100	22	AAK6547	Human bone marrow
C	87	24.6	0.5	100	22	ABS20925	Human genome-deriv
C	88	24.4	0.5	50	22	AAL34109	Human SNP oligonuc
C	89	24.4	0.5	81	22	AAD11241	91-fl primer used
C	90	24.4	0.5	87	24	AAD30431	29 mer oligonucleo
C	91	24.4	0.5	90	24	ABK36835	Human DNA encoding
C	92	24.4	0.5	94	15	AAQ44720	Oligonucleotide us
C	93	24.4	0.5	96	22	ABA17038	Human nervous syst
C	94	24.4	0.5	96	22	ABA08157	Human ovarian and
C	95	24.4	0.5	96	22	AAS29165	Genomic sequence #
C	96	24.4	0.5	96	22	AAI07123	Human reproductive
C	97	24.4	0.5	96	22	AAK73938	Human immune/haema
C	98	24.4	0.5	96	22	AAK89377	Human digestive sy
C	99	24.4	0.5	97	24	ABN63907	Human cancer relat
C	100	24.2	0.5	41	22	AAH73706	Human zinc finger
C	101	24.2	0.5	50	22	AAI28998	Human SNP oligonuc
C	102	24.2	0.5	57	19	AAV17235	SCA2 gene CAG repe
C	103	24.2	0.5	60	24	ABNA3593	Human spliced tran
C	104	24.2	0.5	61	22	AAK32488	Human CB38 exon 1
C	105	24.2	0.5	69	15	AAQ54075	DNA encoding a bom
C	106	24.2	0.5	81	18	AAI76102	Human beta tryptas
C	107	24.2	0.5	81	20	AAK53910	Human beta tryptas
C	108	24.2	0.5	81	21	AAI19475	Human beta tryptas
C	109	24.2	0.5	81	21	AAK33353	Low adenosine anti
C	110	24.2	0.5	87	22	ABA69608	Human foetal liver
C	111	24.2	0.5	87	22	AAI24439	Probe #14372 for g
C	112	24.2	0.5	87	22	AAI49718	Probe #18404 used
C	113	24.2	0.5	87	24	ABSI17852	Human genome-deriv
C	114	24.2	0.5	90	24	ABK17248	Coupled ligation a
C	115	24.2	0.5	90	24	ABK17293	Coupled ligation a
C	116	24.2	0.5	94	22	AAK75496	Codon-optimised HP
C	117	24.2	0.5	94	17	AAI27117	Hepatitis E virus
C	118	24.2	0.5	99	18	AAI96967	Hepatitis E virus
C	119	24.2	0.5	100	22	AAI76819	Codon-optimised HP
C	120	24.2	0.5	100	24	ABL80050	Human ovarian canc
C	121	24	0.5	24	24	ABT03609	Human GIL-3 gene p
C	122	24	0.5	72	17	AAI37259	VDJ joint sequence
C	123	24	0.5	72	18	AAI12554	VDH21 D1 J4 mu c1o
C	124	24	0.5	80	21	ABL50871	Modified erythro
C	125	24	0.5	84	17	AAI33691	Nucleic acid encod
C	126	24	0.5	84	17	AAI33693	Nucleic acid encod
C	127	24	0.5	84	17	AAH42793	Nucleotide sequenc
C	128	24	0.5	87	15	AAQ44722	Oligonucleotide us
C	129	24	0.5	87	20	AAZ28556	Oligonucleotide C2
C	130	24	0.5	91	22	AAI16706	dGMP-specific apta
C	131	24	0.5	97	24	ABL39667	Synthetic Env poly
C	132	24	0.5	100	15	AAO62583	Candida albicans-s
C	133	24	0.5	100	15	AAO62586	Candida albicans-s
C	134	24	0.5	100	18	AAI76186	Human IL4 receptor
C	135	24	0.5	100	21	AAK54015	Human IL-4 recepto
C	136	24	0.5	100	21	AAI19581	Human IL4 receptor
C	137	24	0.5	100	21	AAA33459	Low adenosine anti
C	138	23.8	0.5	45	16	AAO99528	Human Fas ligand g
C	139	23.8	0.5	60	24	ABN34057	Human spliced tran
C	140	23.8	0.5	61	20	AAZ19582	Complement system
C	141	23.8	0.5	65	24	ABN27826	Rat spliced transc
C	142	23.8	0.5	65	24	ABN52293	Mouse spliced tran
C	143	23.8	0.5	73	21	AAI1155	Human secreted pro
C	144	23.8	0.5	81	18	AAI76102	Human beta tryptas
C	145	23.8	0.5	81	18	AAI76102	Human beta tryptas
C	146	23.8	0.5	81	21	AAI19475	Human beta tryptas
C	147	23.8	0.5	81	21	AAA33535	Human beta tryptas
C	148	23.8	0.5	82	19	AAV39759	Low adenosine anti
C	149	23.8	0.5	90	12	AAQ11762	Oligonucleotide SE
C	150	23.8	0.5	90	19	AAV45801	Self-complementary
C	151	23.8	0.5	90	24	ABK36437	Vector PROCOS/3 i
C	152	23.8	0.5	94	19	AAV57539	HCV DNA encoding H
C	153	23.8	0.5	94	21	AAV57539	Human calcium chan
C	154	23.8	0.5	94	21	AAI36682	RSV and PIV fusion
C	155	23.8	0.5	94	22	AAI36682	Human EST H55223
C	155	23.8	0.5	96	21	AAA90615	Oligonucleotide ON

C	99	22	AAH76169	Plasmid PAMV74 co
C	100	17	AAI30902	Primer 32 for 95 k
C	101	22	AAI27311	Human SNP oligonuc
C	102	22	AAI27311	Mouse spliced tran
C	103	22	AAI27311	Trypsin inhibitor
C	104	22	AAI27311	Oligonucleotide se
C	105	22	AAI27311	OST311 related ol
C	106	22	AAI27311	Human DNA encoding
C	107	22	AAI27311	Human breast cell
C	108	22	AAI27311	Human breast cell
C	109	22	AAI27311	Human foetal liver
C	110	22	AAI27311	Human foetal liver
C	111	22	AAI27311	Human foetal liver
C	112	22	AAI27311	Human foetal liver
C	113	22	AAI27311	Human foetal liver
C	114	22	AAI27311	Human foetal liver
C	115	22	AAI27311	Human foetal liver
C	116	22	AAI27311	Human foetal liver
C	117	22	AAI27311	Human foetal liver
C	118	22	AAI27311	Human foetal liver
C	119	22	AAI27311	Human foetal liver
C	120	22	AAI27311	Human foetal liver
C	121	22	AAI27311	Human foetal liver
C	122	22	AAI27311	Human foetal liver
C	123	22	AAI27311	Human foetal liver
C	124	22	AAI27311	Human foetal liver
C	125	22	AAI27311	Human foetal liver
C	126	22	AAI27311	Human foetal liver
C	127	22	AAI27311	Human foetal liver
C	128	22	AAI27311	Human foetal liver
C	129	22	AAI27311	Human foetal liver
C	130	22	AAI27311	Human foetal liver
C	131	22	AAI27311	Human foetal liver
C	132	22	AAI27311	Human foetal liver
C	133	22	AAI27311	Human foetal liver
C	134	22	AAI27311	Human foetal liver
C	135	22	AAI27311	Human foetal liver
C	136	22	AAI27311	Human foetal liver
C	137	22	AAI27311	Human foetal liver
C	138	22	AAI27311	Human foetal liver
C	139	22	AAI27311	Human foetal liver
C	140	22	AAI27311	Human foetal liver
C	141	22	AAI27311	Human foetal liver
C	142	22	AAI27311	Human foetal liver
C	143	22	AAI27311	Human foetal liver
C	144	22	AAI27311	Human foetal liver
C	145	22	AAI27311	Human foetal liver
C	146	22	AAI27311	Human foetal liver
C	147	22	AAI27311	Human foetal liver
C	148	22	AAI27311	Human foetal liver
C	149	22	AAI27311	Human foetal liver
C	150	22	AAI27311	Human foetal liver
C	151	22	AAI27311	Human foetal liver
C	152	22	AAI27311	Human foetal liver
C	153	22	AAI27311	Human foetal liver
C	154	22	AAI27311	Human foetal liver
C	155	22	AAI27311	Human foetal liver

229	23.2	0.5	51	22	AAI78742	Human silent SNP c	302	22.8	0.5	65	24	ABN27731	Rat spliced transc
230	23.2	0.5	51	22	AAI78743	Human silent SNP c	303	22.8	0.5	65	24	ABN27855	Rat spliced transc
231	23.2	0.5	51	22	AAI78743	Human silent SNP c	304	22.8	0.5	65	24	ABN27855	Mouse spliced tran
232	23.2	0.5	72	24	ABL51932	Reverse assembly r	305	22.8	0.5	69	23	ABL53319	Artificial secret
233	23.2	0.5	79	24	ABL50622	OST311N related ol	306	22.8	0.5	76	14	AA042223	PCR primer RNA-76
234	23.2	0.5	81	17	AAI88564	Secretory phosphol	307	22.8	0.5	76	16	AA046142	Fd gene 3' (consta
235	23.2	0.5	82	14	AA057493	Probe (v) used in	308	22.8	0.5	76	16	AA092350	Human secreted pro
236	23.2	0.5	83	19	AAI22179	Oligonucleotide se	309	22.8	0.5	81	21	AAI12506	Oligonucleotide us
237	23.2	0.5	83	19	AAI22179	EBSI gene construc	310	22.8	0.5	82	21	AAI05643	Oligonucleotide SE
238	23.2	0.5	90	22	AAH84353	Human cell death p	311	22.8	0.5	85	21	AAI70571	Oligonucleotide NR
239	23.2	0.5	90	24	ABO90698	M. capsulatus gene	312	22.8	0.5	85	21	AAI70571	VEGF nucleic acid
240	23.2	0.5	90	24	ABO90698	Coupled ligation a	313	22.8	0.5	86	17	AAI88647	Modified erythro
241	23.2	0.5	90	24	ABK17249	Coupled ligation a	314	22.8	0.5	86	21	ABO50869	Human Ig Cmu gene
242	23.2	0.5	91	21	AAA63165	Il reverse gene co	315	22.8	0.5	87	21	AAI60398	Human secreted pro
243	23.2	0.5	91	21	AAA63165	Human secreted pro	316	22.8	0.5	90	20	AAI51378	Oligonucleotide SE
244	23.2	0.5	91	21	AAI11338	Padlock probe alph	317	22.8	0.5	90	20	AAI51378	Human tryptase-I a
245	23.2	0.5	92	16	AAI88439	M13 bacteriophage	318	22.8	0.5	90	21	AAI19478	Human secreted pro
246	23.2	0.5	93	16	AAI88439	Human secreted pro	319	22.8	0.5	90	21	AAI19478	Human tryptase-I a
247	23.2	0.5	94	21	AAO51750	Oligonucleotide pr	320	22.8	0.5	91	22	AAI33356	Human brain expres
248	23.2	0.5	98	14	AAO51750	Pseudomonas aerugi	321	22.8	0.5	91	22	AAI19445	Human bone marrow
249	23.2	0.5	98	23	AAI84771	Human spliced tran	322	22.8	0.5	91	22	AAI19445	Human bone marrow
250	23.2	0.5	60	24	ABN34244	Human secreted pro	323	22.8	0.5	91	22	AAI19445	Probe #20064 used
251	23.2	0.5	60	24	ABN34244	Human secreted pro	324	22.8	0.5	95	22	AAI60637	Human genome-deriv
252	23.2	0.5	65	21	AAI1733	Human secreted pro	325	22.8	0.5	96	22	AAI60637	Probe #10. Uniden
253	23.2	0.5	65	21	AAI1733	Human secreted pro	326	22.8	0.5	96	22	AAI60637	Arabidopsis thalia
254	23.2	0.5	65	24	ABN51286	Mouse spliced tran	327	22.8	0.5	97	22	AAI60637	Human foetal liver
255	23.2	0.5	78	24	ABN51286	PCR primer 6, used	328	22.8	0.5	97	22	AAI60637	Human brain expres
256	23.2	0.5	82	15	AAO44724	Oligonucleotide C2	329	22.8	0.5	97	22	AAI60637	Human bone marrow
257	23.2	0.5	82	15	AAO44724	Oligonucleotide C2	330	22.8	0.5	97	22	AAI60637	Probe #18904 used
258	23.2	0.5	83	22	AAI60633	Probe #6. Uniden	331	22.8	0.5	97	23	AAI60637	Pseudomonas aerugi
259	23.2	0.5	85	19	AAV32418	Homo sapiens clone	332	22.8	0.5	99	21	AAI60637	Human genome-deriv
260	23.2	0.5	85	19	AAV32418	Human cDNA clone B	333	22.8	0.5	99	21	AAI60637	Human secreted pro
261	23.2	0.5	89	19	AAV68382	Human MN gene intr	334	22.8	0.5	41	22	AAI77419	Human immune/haema
262	23.2	0.5	89	21	AAI6569	Human MN intron 4	335	22.6	0.4	41	24	AAI77419	Human zinc finger
263	23.2	0.5	89	21	AAI6569	Codon-optimised HP	336	22.6	0.4	41	24	AAI77419	Human zinc finger
264	23.2	0.5	90	22	AAI75485	Human secreted pro	337	22.6	0.4	60	24	ABN5969	Human spliced tran
265	23.2	0.5	92	21	AAI17348	Human secreted pro	338	22.6	0.4	60	24	ABN5969	Human spliced tran
266	23.2	0.5	93	24	ABK81120	Bacillus clausii g	339	22.6	0.4	60	24	ABN5969	Human spliced tran
267	23.2	0.5	93	24	ABK81120	Human secreted pro	340	22.6	0.4	60	24	ABN5969	Human spliced tran
268	23.2	0.5	94	21	AAI17410	Oligonucleotide HL	341	22.6	0.4	63	17	ABN5969	Timothy grass alle
269	23.2	0.5	94	21	AAI17410	Oligonucleotide HL	342	22.6	0.4	65	24	ABN5969	Rat spliced transc
270	23.2	0.5	95	21	AAI17327	Human secreted pro	343	22.6	0.4	65	24	ABN5969	Mouse spliced tran
271	23.2	0.5	95	21	AAI17327	Human secreted pro	344	22.6	0.4	65	24	ABN5969	Human spliced tran
272	23.2	0.5	97	19	AAH76170	Plasmid pAMVBT4 co	345	22.6	0.4	66	24	ABN5969	Bovine embryonic g
273	23.2	0.5	97	19	AAH76170	Oligonucleotide HL	346	22.6	0.4	73	24	ABN5969	Human foetal liver
274	23.2	0.5	97	19	AAH76170	Oligonucleotide HL	347	22.6	0.4	75	21	AAI05794	ARF promoter nucle
275	23.2	0.5	97	19	AAH76170	Oligonucleotide HL	348	22.6	0.4	78	10	AAI05794	Streptavidin displ
276	23.2	0.5	97	22	AAH74752	Oligonucleotide us	349	22.6	0.4	85	22	AAI05794	DNA encoding poly
277	23.2	0.5	97	22	AAH74752	Oligonucleotide us	350	22.6	0.4	85	22	AAI05794	Human foetal liver
278	23.2	0.5	97	22	AAH74752	Oligonucleotide us	351	22.6	0.4	85	22	AAI05794	Human foetal liver
279	23.2	0.5	97	22	AAH74752	Oligonucleotide us	352	22.6	0.4	85	22	AAI05794	Human brain expres
280	23.2	0.5	97	24	ABA95550	Oligonucleotide HL	353	22.6	0.4	85	22	AAI05794	Human bone marrow
281	23.2	0.5	97	24	ABA95550	Oligonucleotide HL	354	22.6	0.4	85	22	AAI05794	Probe #14815 for g
282	23.2	0.5	97	24	ABA95550	Oligonucleotide HL	355	22.6	0.4	85	22	AAI05794	Probe #19156 used
283	23.2	0.5	99	14	AAO36997	Proinsulin oligonu	356	22.6	0.4	89	22	AAI05794	Human genome-deriv
284	23.2	0.5	99	14	AAO36997	Proinsulin oligonu	357	22.6	0.4	89	22	AAI05794	Human foetal liver
285	23.2	0.5	99	22	AAI60634	Probe #7. Uniden	358	22.6	0.4	89	22	AAI05794	Human secreted pro
286	23.2	0.5	100	22	AAI60634	Human immune/haema	359	22.6	0.4	89	22	AAI05794	Human brain expres
287	22.8	0.5	50	15	AAO69506	Human prepro-oxyto	360	22.6	0.4	89	22	AAI05794	Human bone marrow
288	22.8	0.5	50	15	AAO69506	Human prepro-oxyto	361	22.6	0.4	89	22	AAI05794	Probe #17826 for g
289	22.8	0.5	50	18	AAI63968	Human plasmidogen	362	22.6	0.4	89	22	AAI05794	Probe #25362 used
290	22.8	0.5	50	18	AAI63968	Human plasmidogen	363	22.6	0.4	89	22	AAI05794	Human genome-deriv
291	22.8	0.5	50	20	AAI17256	Test sequence from	364	22.6	0.4	90	18	AAI6105	Human tryptase-I a
292	22.8	0.5	50	20	AAI17256	Test sequence from	365	22.6	0.4	90	18	AAI6105	Human DNA encoding
293	22.8	0.5	50	24	ABK82747	DNA binding molecu	366	22.6	0.4	92	22	ABK82747	Human foetal liver
294	22.8	0.5	50	24	ABK82747	DNA binding molecu	367	22.6	0.4	92	22	ABK82747	Probe #19531 for g
295	22.8	0.5	58	22	AAI89412	Human genetic mark	368	22.6	0.4	92	22	AAI89412	Human brain expres
296	22.8	0.5	60	18	AAI96607	Cyclitol-ubiquinol	369	22.6	0.4	92	22	AAI89412	Human bone marrow
297	22.8	0.5	60	24	ABN40398	Human spliced tran	370	22.6	0.4	92	22	AAI89412	Probe #18141 for g
298	22.8	0.5	62	20	AAI89412	Reverse primer SIN	371	22.6	0.4	92	24	AAI89412	Probe #25937 used
299	22.8	0.5	65	18	AAI89412	Reverse primer SIN	372	22.6	0.4	93	17	AAI89412	Human genome-deriv
300	22.8	0.5	65	20	AAI89412	Reverse primer SIN	373	22.6	0.4	93	21	AAI89412	SEI20K monomer oli
301	22.8	0.5	65	21	AAI89412	Reverse primer SIN	374	22.6	0.4	96	16	AAI89412	Plasmid pPT0317 pr
													Anti-human IL-4 MA



c 521	22	0.4	93	22	ABA39360	Probe #17826 for g	594	21.6	0.4	60	24	ABN42468	Human spliced tran
c 522	22	0.4	93	22	AAK3011	Human brain expres	595	21.6	0.4	60	24	ABN47150	Human spliced tran
c 523	22	0.4	93	22	AAK49188	Human bone marrow	c 596	21.6	0.4	62	21	AAK12223	Human spliced pro
c 524	22	0.4	93	22	AAI26445	Probe #16378 for g	597	21.6	0.4	65	24	ABN52336	Mouse spliced tran
c 525	22	0.4	93	22	AAI55023	Probe #23709 used	598	21.6	0.4	65	24	ABN52316	Mouse spliced tran
c 526	22	0.4	93	24	ABN22756	Human genome-deriv	c 599	21.6	0.4	65	24	ABN56112	Mouse spliced tran
c 527	22	0.4	96	22	AAK89782	Human digestive sy	600	21.6	0.4	65	24	ABN56112	Mouse spliced tran
c 528	22	0.4	96	22	AAK89785	Human digestive sy	c 601	21.6	0.4	65	24	ABN56112	Mouse spliced tran
c 529	22	0.4	97	21	AAK18714	Human secreted pro	602	21.6	0.4	68	20	AAK25127	Mouse spliced tran
c 530	22	0.4	99	21	AAK14673	Nucleotide sequenc	c 603	21.6	0.4	69	20	AAK86081	PCR primer used to
c 531	22	0.4	99	22	AAI54496	Probe #23182 used	604	21.6	0.4	69	20	AAK60513	PCR primer used to
c 532	22	0.4	100	21	AAK11867	Human secreted pro	605	21.6	0.4	70	24	ABK16426	ABK16426
c 533	21.8	0.4	50	22	AAI28314	Human SNP oligonuc	c 606	21.6	0.4	71	20	AAK07015	AAK07015
c 534	21.8	0.4	50	22	AAI31324	Human SNP oligonuc	c 607	21.6	0.4	72	21	AAK69994	AAK69994
c 535	21.8	0.4	51	22	AAI26920	Human SNP oligonuc	c 608	21.6	0.4	72	24	ABN72888	ABN72888
c 536	21.8	0.4	51	22	AAI29554	Human SNP oligonuc	c 609	21.6	0.4	73	21	AAK12579	AAK12579
c 537	21.8	0.4	60	21	AAK29503	Codon altered alka	610	21.6	0.4	75	21	AAK11406	AAK11406
c 538	21.8	0.4	60	24	ABN38638	Human spliced tran	c 611	21.6	0.4	76	22	AAK90518	AAK90518
c 539	21.8	0.4	60	24	ABN40018	Human spliced tran	c 612	21.6	0.4	78	20	AAK04557	AAK04557
c 540	21.8	0.4	60	24	ABN47632	Human spliced tran	c 613	21.6	0.4	81	17	AAK15843	AAK15843
c 541	21.8	0.4	60	24	AAK07154	Human spliced tran	c 614	21.6	0.4	82	21	AAK12514	AAK12514
c 542	21.8	0.4	65	24	ABN31635	Rat spliced trans	615	21.6	0.4	82	23	ABK35319	ABK35319
c 543	21.8	0.4	65	24	ABN54406	Mouse spliced tran	616	21.6	0.4	83	20	AAK19771	AAK19771
c 544	21.8	0.4	65	24	ABN57609	Mouse spliced tran	c 617	21.6	0.4	84	16	AAK22835	AAK22835
c 545	21.8	0.4	65	24	ABN58348	Mouse spliced tran	c 618	21.6	0.4	84	15	AAK21972	AAK21972
c 546	21.8	0.4	65	24	ABN58357	Mouse spliced tran	c 619	21.6	0.4	85	15	AAK074761	AAK074761
c 547	21.8	0.4	67	15	AAK06017	Mouse spliced tran	c 620	21.6	0.4	85	22	ABK74482	ABK74482
c 548	21.8	0.4	68	15	AAV79231	COL2A1 5'-primer (	c 621	21.6	0.4	85	22	ABK39326	ABK39326
c 549	21.8	0.4	69	15	AAK063712	Staphylococcus aur	c 622	21.6	0.4	85	22	AAK22949	AAK22949
c 550	21.8	0.4	70	20	AAK20584	Synthetic 2S seed	c 623	21.6	0.4	85	22	AAK49123	AAK49123
c 551	21.8	0.4	75	17	AAK42304	CTLA-4 variable 11	c 624	21.6	0.4	85	22	AAK54953	AAK54953
c 552	21.8	0.4	75	17	AAK42304	Probe sequence whl	c 625	21.6	0.4	85	24	ABK22708	ABK22708
c 553	21.8	0.4	75	22	AAK13077	HLA-B/HLA-E chimera	c 626	21.6	0.4	87	21	AAK55942	AAK55942
c 554	21.8	0.4	78	22	AAK11626	DNA encoding zinc	627	21.6	0.4	87	24	AAK31516	AAK31516
c 555	21.8	0.4	83	19	AAK22180	Human secreted pro	c 628	21.6	0.4	90	22	AAK07235	AAK07235
c 556	21.8	0.4	83	21	AAK28283	Oligonucleotide se	c 629	21.6	0.4	90	22	ABK76766	ABK76766
c 557	21.8	0.4	84	23	AAK47902	EBST gene constuc	c 630	21.6	0.4	90	24	ABK64112	ABK64112
c 558	21.8	0.4	87	22	ABK471258	Enterococcus faeca	c 631	21.6	0.4	90	24	ABK6780	ABK6780
c 559	21.8	0.4	87	22	ABK471258	Human foetal liver	c 632	21.6	0.4	90	24	ABK82732	ABK82732
c 560	21.8	0.4	87	22	AAK19559	Probe #16022 for g	c 633	21.6	0.4	93	9	AAK81186	AAK81186
c 561	21.8	0.4	87	22	AAK45556	Human brain expres	c 634	21.6	0.4	94	24	ABK91737	ABK91737
c 562	21.8	0.4	87	22	AAI25320	Human bone marrow	c 635	21.6	0.4	94	22	ABK50050	ABK50050
c 563	21.8	0.4	87	22	AAI51436	Probe #15253 for g	c 636	21.6	0.4	95	22	ABK67982	ABK67982
c 564	21.8	0.4	87	24	ABK19826	Probe #20182 used	c 637	21.6	0.4	95	22	ABK67982	ABK67982
c 565	21.8	0.4	89	19	AAK48936	Human genome-deriv	c 638	21.6	0.4	95	22	AAK16375	AAK16375
c 566	21.8	0.4	89	19	AAK48936	HSV-1 latency asso	c 639	21.6	0.4	95	22	AAK42126	AAK42126
c 567	21.8	0.4	89	22	AAK24137	Human brain expres	c 640	21.6	0.4	95	22	AAK42126	AAK42126
c 568	21.8	0.4	89	22	AAK24137	Probe #24821 used	c 641	21.6	0.4	95	24	ABK16170	ABK16170
c 569	21.8	0.4	90	24	ABK36558	Oligonucleotide A8	c 642	21.6	0.4	95	24	ABK11007	ABK11007
c 570	21.8	0.4	91	22	AAK65420	HCV DNA encoding H	c 643	21.6	0.4	97	24	AAK17078	AAK17078
c 571	21.8	0.4	93	12	AAK73258	Oligonucleotide A8	c 644	21.6	0.4	99	21	AAK66736	AAK66736
c 572	21.8	0.4	93	21	AAK29660	VDJ joint sequence	645	21.4	0.4	100	22	AAK46402	AAK46402
c 573	21.8	0.4	95	19	AAK48889	DNA coding for the	646	21.4	0.4	49	22	AAK79263	AAK79263
c 574	21.8	0.4	95	21	AAK48889	HSV-1 latency asso	647	21.4	0.4	55	22	AAK79263	AAK79263
c 575	21.8	0.4	95	21	AAK48889	Anti-human VEGF re	648	21.4	0.4	55	22	AAK79263	AAK79263
c 576	21.8	0.4	95	24	AAK70213	Anti-human FLT-1 a	649	21.4	0.4	60	22	AAK89452	AAK89452
c 577	21.8	0.4	95	24	AAK70213	Oligonucleotide MN	650	21.4	0.4	60	22	AAK31119	AAK31119
c 578	21.8	0.4	97	21	AAK15576	Oligonucleotide 4	c 651	21.4	0.4	60	24	ABK39655	ABK39655
c 579	21.8	0.4	98	21	AAK13297	Human secreted pro	c 652	21.4	0.4	60	24	ABK41305	ABK41305
c 580	21.8	0.4	99	16	AAK26834	Human secreted pro	c 653	21.4	0.4	60	24	ABK43680	ABK43680
c 581	21.8	0.4	99	16	AAK26834	Human gene signatu	c 654	21.4	0.4	60	24	ABK49714	ABK49714
c 582	21.6	0.4	47	20	AAK16201	Human M97-2 secret	c 655	21.4	0.4	60	24	ABK50717	ABK50717
c 583	21.6	0.4	47	21	AAK97539	Serine protease in	656	21.4	0.4	60	24	ABK59239	ABK59239
c 584	21.6	0.4	47	22	AAK167501	Probe used in pro	657	21.4	0.4	60	24	ABK59239	ABK59239
c 585	21.6	0.4	47	23	AAK86666	Serine protease in	658	21.4	0.4	62	14	AAK037291	AAK037291
c 586	21.6	0.4	50	22	AAK34154	CNS disorder-relat	659	21.4	0.4	62	16	AAK095011	AAK095011
c 587	21.6	0.4	51	19	AAK28951	Human SNP oligonuc	660	21.4	0.4	62	20	AAK30476	AAK30476
c 588	21.6	0.4	59	19	AAK40231	Plasmid pAMG21 hfg	661	21.4	0.4	62	20	AAK293616	AAK293616
c 589	21.6	0.4	59	12	AAK97165	Escherichia coli B	662	21.4	0.4	62	21	AAK293616	AAK293616
c 590	21.6	0.4	60	16	AAK094939	Oligonucleotide OD	c 663	21.4	0.4	64	21	AAK13823	AAK13823
c 591	21.6	0.4	60	16	AAK094939	Single chain Fv ju	c 664	21.4	0.4	64	21	ABK32088	ABK32088
c 592	21.6	0.4	60	24	ABN36150	Human spliced tran	c 665	21.4	0.4	65	24	ABK53109	ABK53109
c 593	21.6	0.4	60	24	ABN42265	Human spliced tran	c 666	21.4	0.4	65	24	ABK57332	ABK57332
										66	22	AAK14777	AAK14777





C 813	21.2	0.4	87	21	AA245772	PCR primer 1 used
C 814	21.2	0.4	87	22	AA120156	Human breast cancer
C 815	21.2	0.4	87	22	AAH74733	Probe used identifi
C 816	21.2	0.4	87	24	AA037584	Oligo #1 used in t
C 817	21.2	0.4	89	16	AA023700	Human gene signatu
C 818	21.2	0.4	90	21	AA098463	Human colon cancer
C 819	21.2	0.4	90	22	ABA50328	Human breast cell
C 820	21.2	0.4	90	22	ABA68274	Human foetal liver
C 821	21.2	0.4	90	22	ABA35281	Probe #13747 for g
C 822	21.2	0.4	90	22	AAK16649	Human brain expres
C 823	21.2	0.4	90	22	AAK42413	Human bone marrow
C 824	21.2	0.4	90	22	AA123711	Probe #13104 for g
C 825	21.2	0.4	90	22	AA148490	Probe #17176 used
C 826	21.2	0.4	90	22	AA108826	Probe #8817 used t
C 827	21.2	0.4	90	24	AB516466	Human genome-deriv
C 828	21.2	0.4	90	24	AA172464	18 aa linker const
C 829	21.2	0.4	90	24	ABK36561	HCV DNA encoding H
C 830	21.2	0.4	90	24	ABK36582	HCV DNA encoding H
C 831	21.2	0.4	90	24	ABK36714	Human DNA encoding
C 832	21.2	0.4	90	24	ABK36776	Human DNA encoding
C 833	21.2	0.4	90	24	ABK36912	Human DNA encoding
C 834	21.2	0.4	91	16	AA025335	Human gene signatu
C 835	21.2	0.4	92	23	AB151695	Human GFRalpha4 ex
C 836	21.2	0.4	93	16	AA075907	Primer HP2 for amp
C 837	21.2	0.4	93	21	AA050702	Oligonucleotide SE
C 838	21.2	0.4	96	24	AA146571	Human PAPalpha spe
C 839	21.2	0.4	97	22	ABA48198	Human foetal liver
C 840	21.2	0.4	97	22	ABA66077	Probe #11622 for g
C 841	21.2	0.4	97	22	ABA33156	Human brain expres
C 842	21.2	0.4	97	22	AAK41501	Human bone marrow
C 843	21.2	0.4	97	22	AAK40242	Probe #10939 for g
C 844	21.2	0.4	97	22	AA121006	Probe #14943 used
C 845	21.2	0.4	97	22	AA106257	Probe #6712 used t
C 846	21.2	0.4	97	24	AB514282	Human genome-deriv
C 847	21.2	0.4	98	21	AA553447	Human thioredoxin
C 848	21.2	0.4	100	21	ABA72286	Human foetal liver
C 849	21.2	0.4	100	22	AAK20708	Human brain expres
C 850	21.2	0.4	100	22	AAK20708	Human bone marrow
C 851	21.2	0.4	100	22	AAK46854	Human bone marrow
C 852	21.2	0.4	100	23	AA557709	CDNA #385 encoding
C 853	21.2	0.4	100	23	AB184039	Human ovary cance
C 854	21.2	0.4	100	22	AA096247	Human gene single
C 855	21.2	0.4	100	22	AA096248	Human gene single
C 856	21.2	0.4	29	24	ABK30504	Fluorescent probe
C 857	21.2	0.4	41	17	AA112216	Human cyclin C gen
C 858	21.2	0.4	41	19	AA11837	Human cyclin C PCR
C 859	21.2	0.4	41	21	AA188819	Human zinc finger
C 860	21.2	0.4	41	22	AAH77420	Human OZF gene rel
C 861	21.2	0.4	41	24	AA518907	Human zinc finger
C 862	21.2	0.4	41	24	ABA03752	Human mouse histon
C 863	21.2	0.4	45	21	AA081272	Oligonucleotide en
C 864	21.2	0.4	45	22	AAH24422	Plasmid pRK223-3-1
C 865	21.2	0.4	47	21	AA556041	DNA insert used in
C 866	21.2	0.4	49	21	AA097615	Oligonucleotide #1
C 867	21.2	0.4	50	22	AA088525	Human SNP oligonuc
C 868	21.2	0.4	50	22	AA029113	Human SNP oligonuc
C 869	21.2	0.4	50	22	AA033634	Human clone c92782
C 870	21.2	0.4	51	21	AA076656	Human clone c92782
C 871	21.2	0.4	51	21	AA076657	Human SNP oligonuc
C 872	21.2	0.4	51	21	AA076657	Human clone c92782
C 873	21.2	0.4	51	21	AA076657	Human clone c92782
C 874	21.2	0.4	51	22	AAH79652	Human DNA containi
C 875	21.2	0.4	54	21	AA031088	Human secreted pro
C 876	21.2	0.4	54	22	AA031088	PCR primer used to
C 877	21.2	0.4	56	15	AA067104	Mutagenic primer M
C 878	21.2	0.4	57	20	AA116243	Serine protease in
C 879	21.2	0.4	57	20	AA116244	Serine protease in
C 880	21.2	0.4	57	21	AA097586	Oligonucleotide us
C 881	21.2	0.4	57	21	AA097587	Oligonucleotide us
C 882	21.2	0.4	57	22	AA167547	Serine protease in
C 883	21.2	0.4	57	22	AA167548	Serine protease in
C 884	21.2	0.4	57	22	AA167548	Leu-thiudin/alkali
C 885	21.2	0.4	60	24	ABN33316	Human spliced tran
C 886	21.2	0.4	60	24	ABN33316	Human spliced tran
C 887	21.2	0.4	60	24	ABN33316	Human spliced tran
C 888	21.2	0.4	60	24	ABN33316	Human spliced tran
C 889	21.2	0.4	60	24	ABN33316	Human spliced tran
C 890	21.2	0.4	60	24	ABN33316	Human spliced tran
C 891	21.2	0.4	60	24	ABN33316	Human spliced tran
C 892	21.2	0.4	60	24	ABN33316	Human spliced tran
C 893	21.2	0.4	60	24	ABN33316	Human spliced tran
C 894	21.2	0.4	60	24	ABN33316	Human spliced tran
C 895	21.2	0.4	60	24	ABN33316	Human spliced tran
C 896	21.2	0.4	60	24	ABN33316	Human spliced tran
C 897	21.2	0.4	60	24	ABN33316	Human spliced tran
C 898	21.2	0.4	60	24	ABN33316	Human spliced tran
C 899	21.2	0.4	60	24	ABN33316	Human spliced tran
C 900	21.2	0.4	60	24	ABN33316	Human spliced tran
C 901	21.2	0.4	60	24	ABN33316	Human spliced tran
C 902	21.2	0.4	60	24	ABN33316	Human spliced tran
C 903	21.2	0.4	60	24	ABN33316	Human spliced tran
C 904	21.2	0.4	60	24	ABN33316	Human spliced tran
C 905	21.2	0.4	60	24	ABN33316	Human spliced tran
C 906	21.2	0.4	60	24	ABN33316	Human spliced tran
C 907	21.2	0.4	60	24	ABN33316	Human spliced tran
C 908	21.2	0.4	60	24	ABN33316	Human spliced tran
C 909	21.2	0.4	60	24	ABN33316	Human spliced tran
C 910	21.2	0.4	60	24	ABN33316	Human spliced tran
C 911	21.2	0.4	60	24	ABN33316	Human spliced tran
C 912	21.2	0.4	60	24	ABN33316	Human spliced tran
C 913	21.2	0.4	60	24	ABN33316	Human spliced tran
C 914	21.2	0.4	60	24	ABN33316	Human spliced tran
C 915	21.2	0.4	60	24	ABN33316	Human spliced tran
C 916	21.2	0.4	60	24	ABN33316	Human spliced tran
C 917	21.2	0.4	60	24	ABN33316	Human spliced tran
C 918	21.2	0.4	60	24	ABN33316	Human spliced tran
C 919	21.2	0.4	60	24	ABN33316	Human spliced tran
C 920	21.2	0.4	60	24	ABN33316	Human spliced tran
C 921	21.2	0.4	60	24	ABN33316	Human spliced tran
C 922	21.2	0.4	60	24	ABN33316	Human spliced tran
C 923	21.2	0.4	60	24	ABN33316	Human spliced tran
C 924	21.2	0.4	60	24	ABN33316	Human spliced tran
C 925	21.2	0.4	60	24	ABN33316	Human spliced tran
C 926	21.2	0.4	60	24	ABN33316	Human spliced tran
C 927	21.2	0.4	60	24	ABN33316	Human spliced tran
C 928	21.2	0.4	60	24	ABN33316	Human spliced tran
C 929	21.2	0.4	60	24	ABN33316	Human spliced tran
C 930	21.2	0.4	60	24	ABN33316	Human spliced tran
C 931	21.2	0.4	60	24	ABN33316	Human spliced tran
C 932	21.2	0.4	60	24	ABN33316	Human spliced tran
C 933	21.2	0.4	60	24	ABN33316	Human spliced tran
C 934	21.2	0.4	60	24	ABN33316	Human spliced tran
C 935	21.2	0.4	60	24	ABN33316	Human spliced tran
C 936	21.2	0.4	60	24	ABN33316	Human spliced tran
C 937	21.2	0.4	60	24	ABN33316	Human spliced tran
C 938	21.2	0.4	60	24	ABN33316	Human spliced tran
C 939	21.2	0.4	60	24	ABN33316	Human spliced tran
C 940	21.2	0.4	60	24	ABN33316	Human spliced tran
C 941	21.2	0.4	60	24	ABN33316	Human spliced tran
C 942	21.2	0.4	60	24	ABN33316	Human spliced tran
C 943	21.2	0.4	60	24	ABN33316	Human spliced tran
C 944	21.2	0.4	60	24	ABN33316	Human spliced tran
C 945	21.2	0.4	60	24	ABN33316	Human spliced tran
C 946	21.2	0.4	60	24	ABN33316	Human spliced tran
C 947	21.2	0.4	60	24	ABN33316	Human spliced tran
C 948	21.2	0.4	60	24	ABN33316	Human spliced tran
C 949	21.2	0.4	60	24	ABN33316	Human spliced tran
C 950	21.2	0.4	60	24	ABN33316	Human spliced tran
C 951	21.2	0.4	60	24	ABN33316	Human spliced tran
C 952	21.2	0.4	60	24	ABN33316	Human spliced tran
C 953	21.2	0.4	60	24	ABN33316	Human spliced tran
C 954	21.2	0.4	60	24	ABN33316	Human spliced tran
C 955	21.2	0.4	60	24	ABN33316	Human spliced tran
C 956	21.2	0.4	60	24	ABN33316	Human spliced tran
C 957	21.2	0.4	60	24	ABN33316	Human spliced tran
C 958	21.2	0.4	60	24	ABN33316	Human spliced tran
C 959	21.2	0.4	60	24	ABN33316	Human spliced tran
C 960	21.2	0.4	60	24	ABN33316	Human spliced tran
C 961	21.2	0.4	60	24	ABN33316	Human spliced tran
C 962	21.2	0.4	60	24	ABN33316	Human spliced tran
C 963	21.2	0.4	60	24	ABN33316	Human spliced tran
C 964	21.2	0.4	60	24	ABN33316	Human spliced tran
C 965	21.2	0.4	60	24	ABN33316	Human spliced tran
C 966	21.2	0.4	60	24	ABN33316	Human spliced tran
C 967	21.2	0.4	60	24	ABN33316	Human spliced tran
C 968	21.2	0.4	60	24	ABN33316	Human spliced tran
C 969	21.2	0.4	60	24	ABN33316	Human spliced tran
C 970	21.2	0.4	60	24	ABN33316	Human spliced tran
C 971	21.2	0.4	60	24	ABN33316	Human spliced tran
C 972	21.2	0.4	60	24	ABN33316	Human spliced tran
C 973	21.2	0.4	60	24	ABN33316	Human spliced tran
C 974	21.2	0.4	60	24	ABN33316	Human spliced tran
C 975	21.2	0.4	60	24	ABN33316	Human spliced tran
C 976	21.2	0.4	60	24	ABN33316	Human spliced tran
C 977	21.2	0.4	60	24	ABN33316	Human spliced tran
C 978	21.2	0.4	60	24	ABN33316	Human spliced tran
C 979	21.2	0.4	60	24	ABN33316	Human spliced tran
C 980	21.2	0.4	60	24	ABN33316	Human spliced tran
C 981	21.2	0.4	60	24	ABN33316	Human spliced tran
C 982	21.2	0.4	60	24	ABN33316	Human spliced tran
C 983	21.2	0.4	60	24	ABN33316	Human spliced tran
C 984	21.2	0.4	60	24	ABN33316	Human spliced tran
C 985	21.2	0.4	60	24	ABN33316	Human spliced tran
C 986	21.2	0.4	60	24	ABN33316	Human spliced tran
C 987	21.2	0.4	60	24	ABN33316	Human spliced tran
C 988	21.2	0.4	60	24	ABN33316	Human spliced tran
C 989	21.2	0.4	60	24	ABN33316	Human spliced tran
C 990	21.2	0.4	60	24	ABN33316	Human spliced tran
C 991	21.2	0.4	60	24	ABN33316	Human spliced tran
C 992	21.2	0.4	60	24	ABN33316	Human spliced tran
C 993	21.2	0.4	60	24	ABN33316	Human spliced tran
C 994	21.					



PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
 XX  
 DR WPI; 2002-257383/30.  
 XX  
 PT New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of  
 PT a genome, useful for detecting tissue-, pathology-, and  
 PT developmental-specific genes -  
 XX  
 PS Example 1; SEQ ID 25089; 47pp; English.  
 XX  
 CC The present invention describes oligonucleotide libraries for detecting  
 CC messenger RNAs that populate a (sub-)transcriptome, where the  
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
 CC transcription units that populate a genome. The library comprises  
 CC several oligonucleotides, each capable of hybridizing selectively to a  
 CC set of messenger RNAs transcribed from a given transcription unit of  
 CC the genome, which encodes one or more messenger RNA splice variants.  
 CC The oligonucleotide libraries are useful for detecting mRNAs from a  
 CC biological sample, in expression profiling studies, in qualitatively or  
 CC quantitatively characterizing the corresponding transcriptome, and in  
 CC detecting RNA transcripts and splice variants of human or animal  
 CC transcriptomes. The libraries may also be used as specialised mini  
 CC libraries to detect transcripts of a sub-transcriptome under a  
 CC particular biological or pathological state, and so allowing the  
 CC detection of tissue- and pathology-specific genes such as those genes  
 CC only expressed in specific tissue under a specific pathological  
 CC condition; to detect developmental specific genes; and to detect RNA  
 CC transcripts and splice variants of a transcriptome of a patient suffering  
 CC from a particular disorder. ABN27253 to ABN59589 represent  
 CC oligonucleotide sequences from rats, humans and mice, which are used in  
 CC the exemplification of the present invention.  
 CC N.B. The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WPI  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 CC  
 XX  
 SQ Sequence 65 BP; 15 A; 14 C; 18 G; 18 T; 0 other;  
 Query Match 1.1%; Score 55.4; DB 24; Length 65;  
 Best Local Similarity 90.8%; Pred. No. 0.0066;  
 Matches 59; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 4742 AGTCTTTGCTGACCTCCCTGAGCAAGCAAAATCTCTGAGTTATGCAATGAGCT 4801  
 Db 1 AGTCTTTGCTGACCTCCCTGAGCAAGCAAGCAAGTTCTTGCAGTTATGCAATGAGCG 60  
 Oy 4802 TTAGG 4806  
 Db 61 GTAGG 65  
 DE Human single nucleotide polymorphism (SNP) GLI3 1.  
 DE 18-OCT-2001 (first entry)  
 DE  
 DE Human: resequence; genotype; disease; forensic; paternity testing;  
 KW single nucleotide polymorphism; SNP; ss.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Variation replace(16,C)  
 FT /\*tag= a  
 FT /standard\_name= "single nucleotide polymorphism"  
 XX  
 XX WO200166800-A2.

PD 13-SEP-2001.  
 XX  
 PF 07-MAR-2001; 2001WO-US07268.  
 XX  
 PR 07-MAR-2000; 2000US-0187510.  
 PR 22-MAY-2000; 2000US-0206129.  
 XX  
 PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX  
 PI Cargill M, Ireland JS, Lander ES;  
 XX  
 DR WPI; 2001-522952/57.  
 XX  
 PT Nucleic acid molecules from the human genome which include polymorphic  
 PT sites, useful in methods for predicting the presence, absence or  
 PT severity of a particular phenotype or disorder (e.g. diabetes)  
 PT associated with a particular genotype -  
 XX  
 PS Claim 1; Page 117; 145pp; English.  
 XX  
 CC The invention relates to the identification of nucleic acid molecules  
 CC (AAI29513-AAI31314) from the human genome which include polymorphic sites  
 CC which can predispose individuals to disease. Various genes from a number  
 CC of individuals were resequenced and single nucleotide polymorphisms  
 CC (SNPs) in these genes discovered. The method is useful for predicting the  
 CC presence, absence or severity of a particular phenotype or disorder (e.g.  
 CC diabetes) associated with a particular genotype. The nucleic acids  
 CC containing the polymorphic sites may be useful in forensics and paternity  
 CC testing.  
 CC  
 XX  
 SQ Sequence 31 BP; 5 A; 15 C; 10 G; 1 T; 0 other;  
 Query Match 0.6%; Score 31; DB 22; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-03;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2865 AGGAGGCGCGCGCGCGAGCGCCCTGCGCAAC 2895  
 Db 1 AGGAGGCGCGCGCGCGAGCGCCCTGCGCAAC 31  
 DE Human single nucleotide polymorphism (SNP) GLI3 2.  
 DE 18-OCT-2001 (first entry)  
 DE  
 DE Human: resequence; genotype; disease; forensic; paternity testing;  
 KW single nucleotide polymorphism; SNP; ss.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Variation replace(16,C)  
 FT /\*tag= a  
 FT /standard\_name= "single nucleotide polymorphism"  
 XX  
 XX WO200166800-A2.  
 XX  
 PD 13-SEP-2001.  
 XX  
 PF 07-MAR-2001; 2001WO-US07268.  
 XX  
 PR 07-MAR-2000; 2000US-0187510.  
 PR 22-MAY-2000; 2000US-0206129.  
 XX  
 PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX  
 PI Cargill M, Ireland JS, Lander ES;  
 XX





XX Sequence 88 BP; 30 A; 24 C; 32 G; 2 T; 0 other;  
 SQ Query Match 0.6%; Score 28; DB 22; Length 88;  
 Best Local Similarity 60.5%; Pred. No. 1.1e+04;  
 Matches 46; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 2565 CGTCACTAGTCTGACATGCTCAACAGAGGAGCAGCGCCGACCATCATGCTCGGC 2624  
 DB 3 CGTCACTAGTCTGACATGCTCAACAGAGGAGCAGCGCCGACCATCATGCTCGGC 62

OY 2625 CTACCTGAGAGCGCCG 2640  
 DB 63 CAGCAGGAGCAGCAGC 78

RESULT 10  
 AAS13081  
 ID AAS13081 standard; DNA; 75 BP.  
 AC AAS13081;  
 DT 17-DEC-2001 (first entry)  
 DE DNA encoding zinc finger domain TG-ZFD-040.  
 XX Zinc finger domain; cancer; human; ds.  
 KW Homo sapiens.  
 OS  
 PN WO200160970-A2.  
 PD 23-AUG-2001.  
 PF 17-FEB-2001; 2001WO-KR00244.  
 PR 18-FEB-2000; 2000KR-0007730.  
 PA (TOOL-) TOOLGEN INC.  
 PI Kim J, Kwon Y, Kim H, Ryu EH, Hwang MS;  
 DR P-PSDB; AAU08533.  
 WP: 2001-557644/62.  
 PT Identifying a zinc finger domain for e.g. designing new polypeptides  
 PT that bind to a specific site on a DNA, comprises expressing hybrid  
 PT nucleic acids with a test zinc finger domain in cells -  
 XX Example 49; Page 66; 147bp; English.

XX The invention relates to a method of identifying a zinc finger domain  
 CC that recognises a target site on a DNA. The method comprises expressing  
 CC hybrid nucleic acids with a test zinc finger domain in cells containing a  
 CC reporter construct, where the reporter gene is expressed above a given  
 CC level when a transcription factor recognises a recruitment and a target  
 CC site of a promoter, and not only the recruitment site of the promoter.  
 CC The method is used to: (a) identify a zinc finger domain that recognises  
 CC a target site on a DNA; (b) determine whether a test zinc finger domain  
 CC recognises a target site on a promoter; (c) generate a nucleic acid that  
 CC encodes a chimeric zinc finger protein; and (d) identify DNA sequences  
 CC recognised by zinc finger domains. The method can be used to design novel  
 CC polypeptides that bind to a specific site on a DNA. The method can  
 CC facilitate the customised generation of new polypeptides that can  
 CC regulate the expression of a selected target e.g. a gene required by a  
 CC pathogen can be repressed, a gene required for cancerous growth can be  
 CC repressed, or a gene poorly expressed or encoding a mutated protein can  
 CC be activated and overexpressed. The method can be used in vivo which  
 CC enables identification of polypeptides that bind to a specific site on a  
 CC DNA in the intracellular milieu. The present sequence represents the  
 CC coding sequence of TG-ZFD-040 zinc finger protein which was used in  
 CC the method of the invention.

SQ Sequence 75 BP; 17 A; 25 C; 22 G; 11 T; 0 other;  
 Query Match 0.5%; Score 27.6; DB 22; Length 75;  
 Best Local Similarity 67.2%; Pred. No. 1.2e+04;  
 Matches 39; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 1496 ACTGCCACTGGGAGGCTGCCGCGAGGAGTTCGACCCAGAGAGCTTGACCA 1553  
 DB 5 ACTGCCACTGGGAGGCTGCCGCGAGGAGTTCGACCCAGAGAGCTTGACCA 62

RESULT 11  
 ABK36865  
 ID ABK36865 standard; DNA; 90 BP.  
 AC ABK36865;  
 DT 08-MAY-2002 (first entry)  
 DE Human DNA encoding MAGE-3 segment 2.  
 XX Sayine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmonella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia; ds.  
 XX  
 OS Homo sapiens.  
 PN WO200190197-A1.  
 PD 29-NOV-2001.  
 PF 25-MAY-2001; 2001WO-AU00622.  
 PR 26-MAY-2000; 2000AU-0007761.  
 PA (AUSU) UNIV AUSTRALIAN NAT.  
 PI Thomson SA, Ramsdell IA;  
 DR WP: 2002-147575/19.  
 PT P-PSDB; AAU85045.  
 PT New synthetic polypeptides having several different segments of at  
 PT least one parent polypeptide linked together differently compared to  
 PT the linkage in the parent polypeptide, for inducing immune response  
 PT against a pathogen or cancer -  
 XX Example 3; Fig 27; 364bp; English.

XX The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for  
 CC designing the synthetic polypeptides. The synthetic polypeptides and  
 CC polynucleotides are referred to as a vaccine. The synthetic polypeptide is  
 CC useful for modulating immune responses preferably directed against a  
 CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,  
 CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone  
 CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium) or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present  
 CC sequence encodes a peptide derived from a parent protein used to  
 CC construct a vaccine of the invention.

CC	antigenic i-antigen polypeptide obtained is also useful for treating or preventing i. multifiliis infection in fish. Sequences AAA97036-A97042, CC and AAA97060, AAA97065 and AAA97089 represent i-antigen genes and gene fragments identified in the invention. Sequences AAA97043-A97064 CC (excluding AAA97060) and AAA97071-A97088 represent primers used in the isolation of the i-antigen gene sequences. Sequences AAB25859-B25889 and AAB25893-B25906 represent i-antigen protein and peptide sequences.
XX	Sequence 95 BP; 22 A; 19 C; 33 G; 21 T; 0 other;
Q0	Query Match 0.5%; Score 27.6; DB 17; Length 95; Best Local Similarity 60.8%; Pred. No. 1.3e+04; Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0
QY	3227 GAACACTGCACCTGTCCTGCTCCAGCATCAGCAGAACGTCACCTGGAGTCCC 3286 
Db	82 GAAGTCTCAGTGTATACCCTGTGCTGCTGATATCAACCTGTATAGTGTACGCTACCC 23
QY	3287 TGACCATGACGCT 3300 
Db	22 TGGGAACGACGCT 9
RESULT 13	
ID	AAT27118 standard; DNA; 99 BP.
AC	AAT27118
XX	AAT27118;
XX	06-AUG-1996 (first entry)
DT	Hepatitis E virus (Mexico strain) 406.4-2 antigen DNA.
XX	Hepatitis E virus (Mexico strain) 406.4-2 antigen DNA.
DE	Hepatitis E virus (Mexico strain) 406.4-2 antigen DNA.
XX	HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;
KW	diagnosis; antigen; ds.
XX	Hepatitis E virus Mexico strain.
OS	W09612807-A2.
PN	02-MAY-1996.
XX	23-OCT-1995; 95WO-US13703.
XX	13-OCT-1995; 95US-0542634.
PF	24-OCT-1994; 94US-0327952.
PR	(GENE-) GENELABS TECHNOLOGIES INC.
PA	Fuerst TR, McAttee CP, Yarbough PO, Zhang Y;
XX	WPI; 1996-230608/23.
DR	P-PSDB; AAR96100.
XX	Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as
PT	diagnostic reagents for determining HEV infection and in vaccines
XX	Disclosure; Page 78; 125pp; English.
PS	A DNA sequence (AAT27118) codes for a C-terminal fragment (AAR96100),
CC	406.4-2, of a structural protein (see also AAR96098) of hepatitis
CC	E virus (HEV) Mexico strain. It can be obt. from phase lambda
CC	gcl1 clone 406.3-2, which includes the HEV Mexico ORF-2 (AAT27108)
CC	and ORF-3 (AAT27116). The DNA is inserted into pGEX for expression
CC	of 406.4-2 antigen in E. coli. The antigen can be utilised in
CC	diagnostic assays of HEV when used in conjunction with 62K antigens
CC	(see also AAR96091-92 and AAR96101-04) of HEV capsid protein.
XX	Sequence 99 BP; 12 A; 46 C; 29 G; 12 T; 0 other;
XX	Query Match 0.5%; Score 27.4; DB 17; Length 99;
XX	Best Local Similarity 55.9%; Pred. No. 1.5e+04;
XX	Matches 52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;







XX OS Homo sapiens.  
 XX PN WO200147944-A2.  
 XX PD 05-JUL-2001.  
 XX PF 28-DEC-2000; 2000MO-US35498.  
 XX PR 28-DEC-1999; 99US-0173419.  
 XX PR 27-DEC-2000; 2000US-0173419.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Leach M;  
 XX WPI; 2001-465210/50.  
 XX DR WPI; 2001-465210/50.  
 XX PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 XX PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 XX PT cancer, autoimmune diseases and infections -  
 XX PS Claim 1; Page 1434; 4143pp; English.  
 XX CC The present invention relates to oligonucleotides encoding polymorphic  
 XX CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 XX CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 XX CC histones, kinases, colony stimulating factors, complement related  
 XX CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
 XX CC G-protein coupled receptors and thioesterases. The present sequence is  
 XX CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
 XX CC by them may be used in the prevention, diagnosis and treatment of  
 XX CC diseases associated with inappropriate expression of the proteins listed  
 XX CC above. Disorders that may be prevented, diagnosed and/or treated include  
 XX CC multifactorial diseases with a genetic component, such as autoimmune  
 XX CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 XX CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 XX CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
 XX CC leukemia), diseases of the nervous system and an infection of pathogenic  
 XX CC organisms.  
 XX SQ Sequence 51 BP; 19 A; 11 C; 10 G; 11 T; 0 other;  
 Query Match 0.5%; Score 26.2; DB 22; Length 51;  
 Best Local Similarity 72.3%; Pred. No. 2.2e+04;  
 Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
 QY 1737 AACTTGAACACACTTGATCTCACACTGAGAGAAACCATACG 1783  
 Db 2 AACTGATTGTACATCAGAGAACTCATACGAGAGAAACCTATG 48  
 RESULT 19  
 AAV68382/C  
 ID AAV68382 standard; DNA; 89 BP.  
 XX AC AAV68382;  
 XX DT 10-MAR-1999 (first entry)  
 XX DE Clone #8 fragment identified by CAG repeat analysis method.  
 XX KW CAG repeat; human; genome analysis; medical diagnostic;  
 XX KW nucleic acid analysis; variation assessment; neurological disease;  
 XX KW Huntington's chorea; PCR suppression; SS.  
 XX OS Homo sapiens.  
 XX PN WO9849345-A1.  
 XX PN 05-NOV-1998.  
 XX PD 05-NOV-1998.  
 XX PF 29-APR-1998; 98WO-US08616.

XX XX 29-APR-1997; 97US-0045078.  
 XX PR (UIBO-) UNIV BOSTON.  
 XX PA Smith CL;  
 XX PI Smith CL;  
 XX DR WPI; 1998-594983/50.  
 XX PT Analysing nucleic acid samples - using amplification primers which  
 XX PT contain CAG or CTG tri-nucleotide repeats for differential display  
 XX PT of samples from different sources  
 XX PS Example; Page 32; 44pp; English.  
 XX CC This sequence represents a fragment of a human CAG repeat containing  
 XX CC clone DNA sequence isolated using the method of the invention. The method  
 XX CC is for analysing nucleic acids in a sample, and comprises: (a) providing  
 XX CC a sample containing nucleic acid, a first oligonucleotide primer  
 XX CC comprising a CTG repeat, a second oligonucleotide primer comprising a  
 XX CC CAG repeat and a polymerase and PCR reagents; (b) preparing the nucleic  
 XX CC acid so that it is amplifiable; (c) amplifying the nucleic acid with the  
 XX CC first and second primers; and (d) detecting the amplified product. The  
 XX CC method is used to distinguish between the expression of genes in two or  
 XX CC more biological samples, e.g. body fluids, cells, solid tissue or solid  
 XX CC and liquid foods. It can be used in medical diagnostics, e.g. to  
 XX CC differentiate between normal and diseased tissue or to assess the  
 XX CC variation within monozygotic twin pairs. The method allows the isolation  
 XX CC and analysis of genome subsets containing CAG repeats which are known to  
 XX CC be important in a number of neurological diseases including Huntington's  
 XX CC chorea. The method uses PCR suppression, in which only fragments which  
 XX CC contain a target repeat are efficiently amplified. This allows accurate  
 XX CC identification of differentially expressed genes in various cell types.  
 XX CC Genome complexity is reduced by the new method which targets genomic  
 XX CC subsets containing CAG repeats.  
 XX SQ Sequence 89 BP; 2 A; 23 C; 37 G; 27 T; 0 other;  
 Query Match 0.5%; Score 26.2; DB 19; Length 89;  
 Best Local Similarity 56.3%; Pred. No. 2.7e+04;  
 Matches 49; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
 QY 2598 CAGCAGCCGACACCATCAGCTCGGCTTACCTGAGCAGCCGCTTCAGGATCTC 2657  
 Db 89 CAGCAGCAGCAGCAGCAGCAGCCGCTTCAGGATCTCAGGATCTC 30  
 QY 2658 GCCCTGCTTCTCCAGCCGCCGCTCCAG 2684  
 Db 29 CAGCCGACACGACGACGACGACGAG 3  
 RESULT 20  
 AAX53913  
 ID AAX53913 standard; DNA; 90 BP.  
 XX AC AAX53913;  
 XX DT 05-JUL-1999 (first entry)  
 XX DE Human tryptase-I antisense oligonucleotide fragment.  
 XX KW Antisense oligonucleotide; multiple target; antisense treatment;  
 XX KW impaired respiration; inflammation; lung disease;  
 XX KW pulmonary vasoconstriction; inflammation; allergic rhinitis;  
 XX KW acute asthma; allergy; asthma; impeded respiration;  
 XX KW respiratory distress syndrome; pain; cystic fibrosis;  
 XX KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
 XX KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
 XX KW colon cancer; breast cancer; lung cancer; pancreatic cancer;  
 XX KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
 XX KW prostate cancer; ss.  
 XX OS Synthetic.



RESULT 22  
AAA33356

ID AAA33356 standard; DNA; 90 BP.

AC AAA33356;

DT 28-JUL-2000 (first entry)

DE Low adenosine antisense oligonucleotide SEQ ID NO:1045.

Human: adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; antiinflammatory; antiallergic; antihistaminic; cytosolic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

OS Homo sapiens.

PN WO200009525-A2.

PD 24-FEB-2000.

PF 03-AUG-1999; 99WO-US17712.

PR 03-AUG-1998; 98US-0095212.

PA (UYEC-) UNITV EAST CAROLINA.

PI Nyce JW;

DR WPI: 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary  
PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
PT cancers -

PS Claim 18; Page 396; 1343pp; English.

XX The present invention describes a new composition comprising an  
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which  
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or  
CC inflammation. The ON can have antiinflammatory, antiallergic,  
CC antihistaminic, cytosolic and analgesic activities. The compositions are  
CC useful for the treatment of diseases associated with inflammation,  
CC impaired airways, including lung disease and diseases whose secondary  
CC effects afflict the lungs of a subject. They can be used for treating  
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,  
CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic  
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
CC pulmonary disease (COPD), and cancers such as leukaemia, lymphomas,  
CC carcinomas, and cancers which may metastasize to the lungs, including  
CC breast and prostate cancer. The reduction of the adenosine content of  
CC the ONs reduces side effects. The A-containing ONs break down with the  
CC release of deoxyadenosine which activates adenosine receptors causing  
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the  
CC nucleotide sequences given in the sequence listing from the present  
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last  
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences  
CC differ from the previously named sequences. SEQ ID NO:11 to 1680  
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present  
CC invention. N.B. Sequences given in the disclosure of the present  
CC invention do not match up with their corresponding SEQ ID NO: sequences  
CC given in the sequence listing.

SQ Sequence 90 BP; 0 A; 31 C; 35 G; 9 T; 15 other;

Query Match 0.5%; Score 26.2; DB 21; Length 90;

Best Local Similarity 51.8%; Pred. No. 2.7e+04;

Matches 43; Conservative 9; Mismatches 31; Indels 0; Gaps 0;

QY 2977 GCCCGAGAGGTGACGAGCGGAGACCCACGGCTACGGGCGGCGACACTGCACCG 3036

DB 5 GBCGTGGGCGBGCGCGCGCTGCGCCBGBGCGGCGCGCGCGCGCGCGCG 64

QY 3037 CACGATGCGCTGGGCGCACGGCGT 3059

DB 65 GCTCBGCTCCTGGCGCGCGGAT 87

## RESULT 23

ID AAT78908 standard; cDNA; 66 BP.

AC AAT78908;

DT 09-FEB-1998 (first entry)

DE Poly-glutamine repeat region coding sequence from clone DAN1.

XX Monoclonal antibody; neurodegenerative disease; polyglutamine; TBP;  
KW repeat region; affinity; TBP binding protein; Kennedy disease;  
KW transcription initiation factor; lymphoblastic cell line; schizophrenia;  
KW Huntington's disease; dominant autosomal spinocerebellar ataxia;  
KW X-linked spinocerebellar ataxia; muscular atrophy; familial spastic paraplegia;  
KW dentatorubral-pallidolusial atrophy; bipolar affective disorder;  
KW manic depressive psychosis; ss.

OS Homo sapiens.

PN WO9717445-A1.

PD 15-MAY-1997.

PF 08-NOV-1996; 96WO-FR01773.

PR 10-NOV-1995; 95FR-0013576.

PA (CNRS ) CNRS CENT NAT RECH SCI.

PI (INRM ) INSERM INST NAT SANTE &amp; RECH MEDICALE.

PT Lutz Y, Mandel J, Tora L, Trotter J;

DR WPI: 1997-281034/25.

XX Antibody 1C2 used for treating or preventing neuro-degenerative  
PT diseases - associated with proteins containing long poly:glutamine  
PT repeats, e.g. Huntington's disease

PS Claim 21; Page 43; 69pp; French.

XX The invention relates to a monoclonal antibody (MAB) 1C2 for the  
CC treatment of neurodegenerative diseases associated with the presence  
CC of polyglutamine repeat regions. This MAB is already known for its  
CC affinity to the TBP binding protein (TBP) transcription initiation  
CC factor, especially at the amino acid sequence LERQORQOQOQ found at  
CC the N-terminus of TBP. MAB 1C2 has been shown to have a high affinity  
CC for polyglutamine repeats with a proportional affinity to the number  
CC of glutamine repeats. This affinity has been used to identify genes  
CC encoding proteins containing long polyglutamine repeats which are  
CC implicated in neurodegenerative diseases. A screen of an expression  
CC library, generated from a lymphoblastic cell line from a patient  
CC suffering from spinocerebellar ataxia (SCA), with MAB 1C2 isolated 6 new  
CC sequences (AAT78906-78911) encoding polyglutamine repeats. This  
CC sequence is derived from clone DAN1 isolated from a patient suffering  
CC from dominant autosomal SCA type 7. MAB 1C2, active fragment of it or  
CC nucleic acids encoding it are specifically used to treat Huntington's  
CC disease, SCA types 1-5 or 7, X-linked spinocerebellar muscular atrophy  
CC (Kennedy disease), dentatorubral-pallidolusial atrophy, dominant  
CC autosomal spinocerebellar ataxia, familial spastic paraplegia, bipolar  
CC affective disorder, manic depressive psychoses and schizophrenia.

SQ Sequence 66 BP; 24 A; 22 C; 20 G; 0 U; 0 other;



```

XX 15-JUL-2002 (first entry)
XX Mouse spliced transcript detection oligonucleotide SEQ ID NO:25034.
XX
DE Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN MO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IP01903.
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of
XX a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes.
XX
PS Example 1; SEQ ID 25034; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridizing selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterizing the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN29589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 65 BP; 22 A; 13 C; 16 G; 14 T; 0 other;

```

```

AAD24301
ID AAD24301 standard; DNA; 74 BP.
XX
AC AAD24301;
XX
DT 07-MAR-2002 (first entry)
XX
DE der 22 DNA related to the invention.
XX
KW Genetic deletion; translocation; mutation; conotruncal defect;
XX DiGeorge syndrome; DGS; CHARGE association; Velocardiofacial syndrome;
XX Shprintzen syndrome; cleft palate; der 22; ds.
XX
OS Unidentified.
XX
PN US6303294-B1.
XX
PD 16-OCT-2001.
XX
PF 07-JUN-1995; 95US-0473319.
XX
PR 10-JUL-1992; 92US-0911534.
XX
PR 04-OCT-1991; 91US-0770758.
XX
PR 22-NOV-1993; 93US-0156672.
XX
PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
XX (UYPE-) UNIV PENNSYLVANIA.
XX
PI Emanuel BS, Budarf ML, Driscoll D;
XX
DR WPI; 2002-033211/04.
XX
XX
PT Novel methods to detect genetic changes associated with DiGeorge
XX syndrome, Velocardiofacial syndrome, CHARGE association, conotruncal
XX defect and/or cleft palate are useful for prenatal screening for the
XX diseases.
XX
PS Example 13; Fig 6C; 56pp; English.
XX
CC The invention relates to methods of detecting genetic deletions,
XX translocations and mutations associated with at least one condition
XX selected from the group consisting of DiGeorge syndrome (DGS),
XX CHARGE association, Velocardiofacial (Shprintzen) syndrome (VCF),
XX conotruncal defect and/or cleft palate in a human patient. DGS is
XX linked to chromosomal deletion of chromosome 22. The method
XX involves identifying in a sample DNA if there are less than 2
XX functional copies of chromosome 22q11 and including locus D22S36 to
XX locus BCR12, indicating a genetic deletion or mutation associated
XX with the conditions. The method is useful for diagnosing DGS, VCF,
XX CHARGE association, conotruncal defect and/or cleft palate,
XX particularly in prenatal monitoring. The present sequence is
XX der 22 DNA related to the invention.
XX
SQ Sequence 74 BP; 15 A; 27 C; 17 G; 15 T; 0 other;

```

```

Query Match 0.5%; Score 25.8; DB 24; Length 74;
Best Local Similarity 60.9%; Pred. No. 3.1e+04;
Matches 42; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```

```

OY 2764 CGCCGCTTCACGACGACGAGCGGCTGCCACGCTGCTCAGCCG 2823
DB 4 CTCACCTTCCTCCACCTTCACGAGAGGCGCTGCAGGCACTCATCTCAGAGTCC 63

```

```

OY 2824 GCCCAGCAG 2832
DB 64 TCCAAAGCTG 72

```

```

RESULT 28
AAC21832
ID AAC21832 standard; cDNA; 84 BP.
XX
AC AAC21832;

```

```

XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 25907.
XX DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-0200610.
XX PR 26-FEB-1999; 99US-0122487.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX CC New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX CC obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX CC diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS Claim 1; SEQ ID 25907; 71pp + CD-ROM; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors.
XX SQ Sequence 84 BP; 15 A; 18 C; 23 G; 28 T; 0 other;
XX
Query Match 0.5%; Score 25.8; DB 21; Length 84;
Best Local Similarity 67.9%; Pred. No. 3.2e+04;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 4838 GGAGTTGAAGAGATTAACTGACTTTGTTGGCTGTTTTTTTGTCTGTAT 4890
DB 16 GGAGAGGAAGAATTTTACAGGCTTTTATTGCGCGTTATTTTCTGTGTCT 68
RESULT 29
AAA36669
ID AAA36669 standard; DNA; 97 BP.
AC AAA36669;
XX DT 01-AUG-2000 (first entry)
XX DE RSV and PIV fusion protein F-RSV/HN-PIV3 oligonucleotide SEQ ID NO:28.
XX DE RSV protein; heterochimeric; immunogenic; infection; RSV; PIV; MV;
XX KW Muv; respiratory syncytial virus; parainfluenza virus; mumps virus;
XX KW measles virus; glycoprotein; vaccine; immunisation; humanised; PCR;
XX KW respiratory disorder; paramyxoviridae viral antigen; ss.
XX OS Respiratory syncytial virus.
XX OS Parainfluenza virus.

```

```

XX WO200018929-A2.
XX PN 06-APR-2000.
XX PD 20-SEP-1999; 99WO-EP07004.
XX PF 25-SEP-1998; 98GB-0020931.
XX PR 24-MAR-1999; 99GB-0006868.
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Bollen A, Howard S;
XX DR WPI; 2000-293163/25.
XX CC Expressing viral (e.g. parainfluenza virus and measles virus)
XX CC heterochimeric proteins comprising fragments of in Chinese Hamster
XX CC Ovary cells, useful as vaccines for treating a human or animal
XX CC susceptible to paramyxoviridae viral infections -
XX PS Example 2; Fig 28; 127pp; English.
XX CC The present invention describes a process for expressing, in Chinese
XX CC Hamster Ovary (CHO) cells, a heterochimeric protein or its immunogenic
XX CC derivative comprising immunogenic fragments of respiratory syncytial
XX CC virus (RSV), parainfluenza virus type I (PIV1), PIV2, PIV3, measles
XX CC virus (MV) or mumps virus (MuV) fusion and attachment glycoproteins.
XX CC A vaccine comprising a protein as described above, can be used for
XX CC treating a human or animal susceptible to paramyxoviridae viral
XX CC infections. The heterochimeric proteins or their immunogenic derivatives
XX CC are used in the manufacture of medicaments for use in the treatment of
XX CC respiratory disorders. AAA36653 to AAA36695 represent oligonucleotides
XX CC used in the generation of PCR fragments as part of the construction of
XX CC an RSV/PIV3 fusion protein nucleotide sequence, and AAA36696 to AAA36699
XX CC represent the RSV/PIV3 fusion protein nucleotide sequences produced in
XX CC an example from the present invention. AAA36700 to AAA36739 represent
XX CC oligonucleotides used in the generation of PCR fragments as part of the
XX CC construction of an Muv/MV fusion protein nucleotide sequence, and
XX CC AAA36740 to AAA36743 represent the nucleotide sequences produced in an
XX CC example from the present invention.
XX SQ Sequence 97 BP; 27 A; 35 C; 18 G; 17 T; 0 other;
XX
Query Match 0.5%; Score 25.8; DB 21; Length 97;
Best Local Similarity 67.9%; Pred. No. 3.4e+04;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 3098 GCCTGGCCCTCCCTCGTGTGCGCGCTTCAGCAGCCTCAGCAGTGCAGAACCC 3150
DB 18 GCCTGACCCCTCCCGAGGCTGACCTCTGCAACATCGACATCTTCAACCCC 70
RESULT 30
ABN36458
ID ABN36458 standard; DNA; 60 BP.
XX AC ABN36458;
XX DT 15-JUL-2002 (first entry)
XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:9206.
XX DE Human; mouse; rat; splice transcript; detection; RNA transcript;
XX KW splice variant; transcriptome; oligonucleotide library; ss.
XX OS Homo sapiens.
XX OS WO200210449-A2.
XX PN 07-FEB-2002.
XX PD 20-JUL-2001; 2001WO-IB01903.
XX PF

```





XX	Key	Location/Qualifiers
FFH	mat_peptide	1..92
FT		/*tag= a
FT		/transl_except= (pos:91..92, aa:Ile)
FT		
XX		
XX		
PN	W09739023-A1.	
XX		
PD	23-OCT-1997.	
XX		
PPF	04-APR-1997;	97WO-S500574.
XX		
PR	23-SEP-1996;	96SE-0003469.
PPR	12-APR-1996;	96SE-0001422.
XX		
XX	(ASTR ) ASTRA AB.	
XX		
PPI	Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;	
XX		
DDR	WPI; 1997-526397/48.	
XX	P-PSDB; AAW45013.	
XX		
PPT	Nucleic acids encoding cysteine- or methionine-containing peptide(s)	
PPT	which have immuno-stimulatory or immunosuppressive activity - can be	
PPT	used to treat, e.g. cancers, infection, auto-immune disease or	
PPT	transplant rejection	
XX		
XX	Claim 14; Page 158; 183pp; English.	
XX		
CCC	The present sequence encodes an immunosuppressing or	
CCC	immunostimulatory peptide. An immunosuppressant can be used to	
CCC	treat transplant rejection or autoimmune disease, e.g. rheumatoid	
CCC	arthritis, systemic lupus erythematosus, Sjogren's syndrome,	
CCC	sceleroderma, mixed connective tissue disease, dermatomyositis,	
CCC	polymyositis, Reiter's syndrome, Behcet's disease, type I diabetes,	
CCC	Hashimoto's thyroiditis, Graves' disease, multiple sclerosis,	
CCC	myasthenia gravis, encephalomyelitis, pemphigus vulgaris, vegetans	
CCC	or foliaceus, Seneear-Usher syndrome or Brazilian pemphigus. An	
CCC	immunostimulator can be used to treat conditions such as cancer or	
CCC	infection.	
XX		
XX	Sequence 92 BP; 4 A; 31 C; 33 G; 24 T; 0 other;	
XX		
XX	Query Match	0.5%; Score 25.6; DB 18; Length 92;
XX	Best Local Similarity	62.5%; Pred. No. 3.7e+04;
XX	Matches	40; Conservative 0; Mismatches 24; Indels 0; Caps 0;
XX		
2QY	2580 CATGCTCAACAGAGGACAGCAGCCAGCACCACATCAGCTGGGCTTACCTGAGCAGCCG	2639
Db		
XX	69 CCTGCAAGGACAGAGGGGACAGAGCCACGACGATGCTGGCGGACACCGAGCCG	10
XX		
2QY	2640 CCGC 2643	
Db		
XX	9 CAGC 6	
XX		
XX	RESULT 33	
XX	RAV00492/C	
XX	ID AAV00492 standard; DNA; 93 BP.	
XX	AAV00492;	
XX		
DT	27-APR-1998 (first entry)	
XX		
DE	DNA for immunomodulatory peptide D22020AX.	
XX		
XX	Immunomodulator; immunosuppressant; immunostimulator; treatment;	
KW	transplant rejection; autoimmune disease; cancer; infection; ss.	
XX		
XX	Synthetic.	
XX		
XX	Key	Location/Qualifiers
FFH	mat_peptide	1..93
FT		

FT		/ *tag= a	
PN	W09739023-A1.		
XX			
PD	23-OCT-1997.		
XX			
PF	04-APR-1997; 97WO-S000574.		
XX			
PR	23-SEP-1996; 96SE-0003469.		
PR	12-APR-1996; 96SE-0001422.		
XX			
PA	(ASTR ) ASTRA AB.		
XX			
PI	Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;		
XX			
WI	WPI; 1997-526397/48.		
P-	P-PSDB; AAW36997.		
XX			
Nucleic acids encoding cysteine- or methionine-containing peptide(s) which have immunostimulatory or immunosuppressive activity - can be used to treat, e.g. cancers, infection, auto:immune disease or transplant rejection			
Claim 36; Page 167; 183pp; English.			
The present sequence encodes an immunosuppressing or immunostimulatory peptide. An immunosuppressant can be used to treat transplant rejection or autoimmune disease, e.g. rheumatoid arthritis, systemic lupus erythematosus, Sjogren's syndrome, scleroderma, mixed connective tissue disease, dermatomyositis, polymyositis, Reiter's syndrome, Behcet's disease, type I diabetes, Hashimoto's thyroiditis, Graves' disease, multiple sclerosis, myasthenia gravis, encephalomyelitis, pemphigus vulgaris, vegetans or foliaceus, Seneear-Usher syndrome or Brazilian pemphigus. An immunostimulator can be used to treat conditions such as cancer or infection.			
Sequence 93 BP; 6 A; 28 C; 34 G; 25 T; 0 other;			
Query Match 0.5%; Score 25.6; DB 18; Length 93;			
Best Local Similarity 62.5%; Pred. No. 3.7e+04;			
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps			
QY 2580 CATGCTCAACAGAGGGACAGCAGCGCCAGCACCATCAGTCGGCTACCTGAGCAGCG 263			
Ddb 69 CCTGCAGGACAGAGGGCGACAGGCCACACAGCAGATGCCCCGGGACACCAGCAGCG 10			
QY 2640 CGCG 2643			
Ddb 9 CAGC 6			
RESULT 34			
ID AAAV05448/C			
ID AAV05448 standard; DNA; 93 BP.			
AAV05448;			
27-APR-1998 (first entry)			
DNA for immunomodulatory peptide D7208.			
Immunomodulator; immunosuppressant; immunostimulator; treatment; transplant rejection; autoimmune disease; cancer; infection; ss.			
Synthetic.			
Key Location/Qualifiers			
mat_peptide 1..93			
FT / *tag= a			
XX W09739023-A1.			

Copied from PCI US2003092 on 04-03-2004



PA	(ASTR ) ASTRA AB.
XX	Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;
PI	
XX	
DR	WPI; 1997-526397/48.
DR	P-PSDB; AAW45017.
XX	
PT	Nucleic acids encoding cysteine- or methionine-containing peptide(s)
PT	which have immuno:stimulatory or immunosuppressive activity - can be
PT	treated to treat, e.g. cancers, infection, auto-immune disease or
PT	transplant rejection
XX	
PS	Claim 14; Page 159; 183pp; English.
XX	
CC	The present sequence encodes an immunosuppressing or
CC	immunostimulatory peptide. An immunosuppressant can be used to
CC	treat transplant rejection or autoimmune disease, e.g. rheumatoid
CC	arthritis, systemic lupus erythematosus, Sjogren's syndrome,
CC	scleroderma, mixed connective tissue disease, dermatomyositis,
CC	polymyositis, Reiter's syndrome, Behcet's disease, type I diabetes,
CC	Hashimoto's thyroiditis, Graves' disease, multiple sclerosis,
CC	myasthenia gravis, encephalomyelitis, phemphigus vulgaris, vegetans
CC	or foliaceus, Seneat-Usher syndrome or Brazilian pemphigus. An
CC	immunostimulator can be used to treat conditions such as cancer or
CC	infection.
XX	
SQ	Sequence 93 BP; 6 A; 30 C; 32 G; 25 T; 0 other;
Query Match	0.5%; Score 25.6; DB 18; Length 93;
Best Local Similarity	62.5%; Pred. No. 3.7e+04;
Matches	40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY	2580 CATGCTCAACAGAGGGACAGCGGCCACCATCAGTCGCCTACCTGAGCAGCGC 2639   69 CTGTGAAGGCACAGACAGAGCGGCAGAGCCACAGCATGCCCGGGACACAGCAGCGC 10
Ddb	
QY	2640 CGCG 2643       9 CAGC 6
Dd	
RESULT 37	
ID	AAV05459/c
XX	AAV05459 standard; DNA; 93 BP.
XX	
XX	AAV05459;
DE	
XX	27-APR-1998 (first entry)
XX	
XX	DNA for immunomodulatory peptide D22184A.
XX	
XX	Immunomodulator; immunosuppressant; immunostimulator; treatment;
KW	transplant rejection; autoimmune disease; cancer; infection; ss.
OS	Synthetic.
FT	
Key	Location/Qualifiers
mat_peptide	1..93
/tag=	a
WO9739023-A1.	
PD	
23-OCT-1997.	
04-APR-1997;	97WO-S000574.
23-SEP-1996;	96SE-0003469.
12-APR-1996;	96SE-0001422.
(ASTR ) ASTRA AB.	
Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;	

DR	WPI; 1997-526397/48.
DR	P-PSDB; AAW45039.
XX	
PT	Nucleic acids encoding cysteine- or methionine-containing peptide(s)
PT	which have immuno:stimulatory or immunosuppressive activity - can be
PT	treated to treat, e.g. cancers, infection, auto-immune disease or
PT	transplant rejection
XX	
PS	Claim 22; Page 163; 183pp; English.
XX	
CC	The present sequence encodes an immunosuppressing or
CC	immunostimulatory peptide. An immunosuppressant can be used to
CC	treat transplant rejection or autoimmune disease, e.g. rheumatoid
CC	arthritis, systemic lupus erythematosus, Sjogren's syndrome,
CC	scleroderma, mixed connective tissue disease, dermatomyositis,
CC	polymyositis, Reiter's syndrome, Behcet's disease, type I diabetes,
CC	Hashimoto's thyroiditis, Graves' disease, multiple sclerosis,
CC	myasthenia gravis, encephalomyelitis, phemphigus vulgaris, vegetans
CC	or foliaceus, Seneat-Usher syndrome or Brazilian pemphigus. An
CC	immunostimulator can be used to treat conditions such as cancer or
CC	infection.
XX	
SQ	Sequence 93 BP; 6 A; 30 C; 34 G; 23 T; 0 other;
Query Match	0.5%; Score 25.6; DB 18; Length 93;
Best Local Similarity	62.5%; Pred. No. 3.7e+04;
Matches	40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY	2580 CATGCTCAACAGAGGGACAGCGGCCACCATCAGTCGCCTACCTGAGCAGCGC 2639   69 CCTGTCAAGGCACAGACAGAGCGGCAGAGCCACAGCATGCCCGGGACACAGCAGCGC 10
Ddb	
QY	2640 CGCG 2643       9 CAGC 6
Dd	
RESULT 38	
AAV05456/c	
ID	AAV05456 standard; DNA; 96 BP.
XX	
XX	AAV05456;
XX	
DT	
XX	27-APR-1998 (first entry)
DE	
XX	DNA for immunomodulatory peptide D22045AX.
XX	
KW	Immunomodulator; immunosuppressant; immunostimulator; treatment;
KW	transplant rejection; autoimmune disease; cancer; infection; ss.
OS	Synthetic.
XX	
Key	Location/Qualifiers
mat_peptide	1..96
/tag=	a
WO9739023-A1.	
PN	
23-OCT-1997.	
04-APR-1997;	97WO-S000574.
23-SEP-1996;	96SE-0003469.
12-APR-1996;	96SE-0001422.
(ASTR ) ASTRA AB.	
Bergstrand H, Eriksson T, Lindvall M, Saernstrand B;	
WPI; 1997-526397/48.	
P-PSDB; AAW45015.	
XX	
XX	Nucleic acids encoding cysteine- or methionine-containing peptide(s)

PT which have immuno:stimulatory or immunosuppressive activity - can be  
 PT used to treat, e.g. cancers, infection, auto:immune disease or  
 PT transplant rejection

XX Claim 14; Page 159; 183pp; English.

CC The present sequence encodes an immunosuppressing or  
 CC immunostimulatory peptide. An immunosuppressant can be used to  
 CC treat transplant rejection or autoimmune disease, e.g. rheumatoid  
 CC arthritis, systemic lupus erythematosus, Sjogren's syndrome,  
 CC scleroderma, mixed connective tissue disease, dermatomyositis,  
 CC polymyositis, Reiter's syndrome, Behcet's disease, type I diabetes,  
 CC Hashimoto's thyroiditis, Graves' disease, multiple sclerosis,  
 CC myasthenia gravis, encephalomyelitis, pemphigus vulgaris, vegetans  
 CC or foliaceus, Seneear-Usher syndrome or Brazilian pemphigus. An  
 CC immunostimulator can be used to treat conditions such as cancer or  
 CC infection.

XX Sequence 96 BP; 5 A; 30 C; 35 G; 26 T; 0 other;

Query Match 0.5%; Score 25.6; DB 18; Length 96;

Best Local Similarity 62.5%; Pred. No. 3.8e+04;

Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 2580 CATGCTCAACAGAGGACAGAGCGCCAGCACCATCATCGCTGAGCAGCGC 2639

DB 69 CTTGAAGGCGACAGAGGCGAGAGAGCCAGCAGCATGCGCGGACACAGCAGCGC 10

QY 2640 CCGC 2643

DB 9 CAGC 6

RESULT 39

ABK43215

ID ABK43215 standard; DNA; 96 BP.

XX ABK43215;

XX 05-JUN-2002 (first entry)

DE Recombinant adenovirus fiber associated primer 418.

XX Modified virus; adenovirus; cytostatic; gene therapy; tumour cell;  
 XX proliferating cell; cancer; vascular disease; inflammatory disease;  
 XX infectious disease; human immunodeficiency virus; HIV; PCR; primer;  
 XX ss.

XX Synthetic.

XX WO200208263-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-GB03252.

XX 19-JUL-2000; 2000GB-0017720.

XX (GOTA-) GOT-A-GENE AB.

XX (GARD/) GARDNER R.

XX Lindholm L, Nord AK, Boulanger PA;

XX WPI; 2002-217049/27.

XX Novel modified virus comprising non-native polypeptides with stable  
 PT conformation and having framework moieties containing binding moieties  
 PT which confer upon the virus, an altered tropism, useful in gene therapy

XX Example 1; Page 46; 163pp; English.

XX The invention describes a modified virus comprising non-native

CC polypeptides which has framework moieties each containing binding  
 CC moieties, where the virus has altered tropism conferred by the binding  
 CC moieties. The polypeptides can be expressed in the cytoplasm and nucleus  
 CC of mammalian host cell in conformation which is maintained in absence of  
 CC ligands for the binding moieties, where the conformation allows the  
 CC binding moiety subsequently to bind with the ligand. The modified virus  
 CC is useful in therapy for the preparation of a medicament for treating  
 CC tumour cells, cancer, proliferating cells, vascular diseases,  
 CC inflammatory diseases and infectious diseases such as Human  
 CC immunodeficiency virus (HIV). The altered tropisms allow the virus to be  
 CC used in treatment of disease in human or animal subjects, either by in  
 CC vivo treatment of, or ex vivo treatment of cells of, the subject  
 CC requiring treatment. The problems associated with the expression of  
 CC functional non-native viral components in the nucleus and cytosol of  
 CC host cells is solved by using the modified virus for the purpose. This  
 CC sequence represents a primer associated with the creation of the  
 CC modified vector, described in the invention.

XX Sequence 96 BP; 8 A; 14 C; 10 G; 10 T; 54 other;

Query Match 0.5%; Score 25.6; DB 24; Length 96;

Best Local Similarity 26.6%; Pred. No. 3.8e+04;

Matches 25; Conservative 17; Mismatches 52; Indels 0; Gaps 0;

QY 2596 GACAGCAGCGCCAGCACCATCATCGCTGAGCAGCGCGCTCTCAGGAGTC 2655

DB 2 GACTGCACCACTGGACCTGCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNS 61

QY 2656 TCGCCCTGCTTCTCCAGCGCCGCTCCAGCGAGG 2689

DB 62 NNSNNNS 95

RESULT 40

ABA72331/c

ID ABA72331 standard; DNA; 98 BP.

XX ABA72331;

XX 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #20636.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -

XX Claim 4; SEQ ID NO 20636; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 98 BP; 18 A; 18 C; 27 G; 35 T; 0 other;

Query Match 0.5%; Score 25.6; DB 22; Length 98;  
Best Local Similarity 62.5%; Pred. No. 3.8e+04;  
Matches 40; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1861 AATGGAACCATATGTGTGCAAAATCCAGGCTGCACTAAGCGTTACACAGACCCCAAGC 1920  
|| ||||| ||| ||||| ||| ||| ||| ||| |||||  
Db 95 AAACAGAAATCACAGATGTGCCAAACTCTGGGCATCTCACAGGGAGAGACAATGCCAAGC 36

QY 1921 TCCC 1924

Db |||  
35 TCAC 32

Search completed: December 11, 2002, 02:41:53  
Job time : 664 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2002, 19:50:51 ; Search time 638 Seconds  
(without alignments)  
17843.027 Million cell updates/sec

Title: US-09-910-185-3  
Perfect score: 5055  
Sequence: 1 cgtactacgtggcatttt.....acccctcttttaaaaaaaaaa 5055

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
N_Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*		
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5055	100.0	5055	24	ABK84036
2	5055	100.0	5055	24	Human cDNA differe
3	3696.6	73.1	5113	24	Human Gli3 coding
4	366	7.2	557	24	Murine Gli3 coding
5	366	7.2	557	24	Oligonucleotide fo
6	334.8	6.6	3635	24	Oligonucleotide fo
7	329.4	6.5	3326	22	Murine Gli1 coding
8	329.4	6.5	3600	22	Human Gli-1 nucleo
9	329.4	6.5	3600	24	Human Cubitus inte
					Human Gli1 coding

10	329.4	6.5	3600	24	AAL45543
11	329.4	6.5	3600	24	AAL45544
12	329.4	6.5	3600	24	AAL45545
13	329.4	6.5	3600	24	ABK30501
14	310.6	6.1	557	24	ABQ15578
15	310.6	6.1	557	24	ABQ15579
16	280.2	5.5	491	22	ABA59457
17	280.2	5.5	491	22	ABA59458
18	280.2	5.5	491	22	ABA59459
19	256	5.1	4853	23	ABSO3987
20	256	5.1	4853	23	ABSO3988
21	204.6	4.0	462	24	ABS08520
22	187.4	3.7	189	24	ABS08521
23	185	3.7	185	24	ABS08522
24	174.2	3.4	466	22	AAH99414
25	171	3.4	2669	23	ABL08099
26	167.8	3.3	470	22	ABA53823
27	167.8	3.3	470	22	AAI33460
28	154	3.0	154	22	ABA71965
29	154	3.0	154	22	AAI52336
30	154	3.0	154	24	ABS20784
31	124.4	2.5	10559	23	ABLO3986
32	124.4	2.4	124	22	ABA66401
33	122	2.4	124	22	AAI46596
34	113.4	2.2	6169	22	AAS46369
35	113.4	2.2	6169	24	ABN80096
36	111.2	2.2	6169	22	AAS46370
37	111.2	2.2	6169	24	ABN80097
38	110.4	2.2	1230	23	ABLO6813
39	109	2.2	3672	23	ABLO6812
40	102	2.0	2947	24	ABI92239
41	99.4	2.0	14486	22	AAS27844
42	93	1.8	2947	23	ABL02109
43	90.8	1.8	3138	21	AAAG2683
44	90.8	1.8	3138	24	ABN95218
45	88.2	1.7	797	24	ABK34450

ALIGNMENTS

RESULT 1	
ABK84036	
ID	ABK84036 standard; cDNA: 5055 BP.
XX	
AC	ABK84036;
XX	
DT	14-AUG-2002 (first entry)
XX	
DE	Human cDNA differentially expressed in granulocytic cells #607.
XX	
KW	Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW	viral infection; parasitic infection; protozoal infection;
KW	fungal infection; sterile inflammatory disease; psoriasis;
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW	adult respiratory distress syndrome; inflammatory bowel disease;
KW	Crohn's disease; ulcerative colitis; periodontal dysery;
KW	granulocyte activation; chronic inflammation; allergy.
OS	Homo sapiens.
XX	
PN	WO200228999-A2.
XX	
PD	11-APR-2002.
XX	
PF	03-OCT-2001; 2001WO-US30821.
XX	
PR	03-OCT-2000; 2000US-237189P.
XX	
PA	(GENE-) GENE LOGIC INC.
XX	
PI	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;







QY 3601 CCGCTGTGCCGAGACTCCGCGCTTTGGGTTCTGCAAGGCATGCTCGTCCACCGCGAG 3660  
 Db 3601 CCGCTGTGCCGAGACTCCGCGCTTTGGGTTCTGCAAGGCATGCTCGTCCACCGCGAG 3660  
 QY 3661 AACCCCTTGAGGAGCGGCGCTGCTGGGGGTATCAGACCCCTCGGGAGAACAGCAACCC 3720  
 Db 3661 AACCCCTTGAGGAGCGGCGCTGCTGGGGGTATCAGACCCCTCGGGAGAACAGCAACCC 3720  
 QY 3721 TAGCGTGGCCAGAGCACTTGATGCTCCACACAGCGCCCGGAAGTGGCACCACTGGAAC 3780  
 Db 3721 TAGCGTGGCCAGAGCACTTGATGCTCCACACAGCGCCCGGAAGTGGCACCACTGGAAC 3780  
 QY 3781 GCCTTCCATGAACAGCCCTGTAAGGCCCGCAGTATGGAACCTGCTCAACAGGCAGCCA 3840  
 Db 3781 GCCTTCCATGAACAGCCCTGTAAGGCCCGCAGTATGGAACCTGCTCAACAGGCAGCCA 3840  
 QY 3841 GTGCCCCCTGGTCACTCGACGCTGCTGTGGTGGCGGCTGCGGCTGCGGCTCAAGCTGAAG 3900  
 Db 3841 GTGCCCCCTGGTCACTCGACGCTGCTGTGGTGGCGGCTGCGGCTGCGGCTCAAGCTGAAG 3900  
 QY 3901 AGCACCCCTGCAAGGAGCGGGGCGGCGAGCTGAATTTGCGGCTGCGGCTGCGGCTCAAGCTGAAG 3960  
 Db 3901 AGCACCCCTGCAAGGAGCGGGGCGGCGAGCTGAATTTGCGGCTGCGGCTGCGGCTCAAGCTGAAG 3960  
 QY 3961 GAGTCAGCTGGCAGCATGCTGAATGTCATGCAAGCAACAGGACCCAGTGGGAGCGGCTAC 4020  
 Db 3961 GAGTCAGCTGGCAGCATGCTGAATGTCATGCAAGCAACAGGACCCAGTGGGAGCGGCTAC 4020  
 QY 4021 CTGGCTCACCAGCTCCTCGGCGACAGCATGCAAGCAACAGGACCCAGTGGGAGCGGCTAC 4080  
 Db 4021 CTGGCTCACCAGCTCCTCGGCGACAGCATGCAAGCAACAGGACCCAGTGGGAGCGGCTAC 4080  
 QY 4081 CAGATGCTTGGGCGAGATTAGTGTCTCTCACACATCAACATCTACCAAGGGCCAGAGAGC 4140  
 Db 4081 CAGATGCTTGGGCGAGATTAGTGTCTCTCACACATCAACATCTACCAAGGGCCAGAGAGC 4140  
 QY 4141 TGCTGCGAGGGGCTCAGGCGATGCGGCGAGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4200  
 Db 4141 TGCTGCGAGGGGCTCAGGCGATGCGGCGAGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4200  
 QY 4201 TACAGCCATGTCAGCTTGGGGGCGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4260  
 Db 4201 TACAGCCATGTCAGCTTGGGGGCGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4260  
 QY 4261 GCTCTGAGTCAAGGAGCGGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4320  
 Db 4261 GCTCTGAGTCAAGGAGCGGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4320  
 QY 4321 ATGAGATGAAGGGGAGCGGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4380  
 Db 4321 ATGAGATGAAGGGGAGCGGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4380  
 QY 4381 TTCTATGACCAAAACCGTGGGCTTCAAGTCAAGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4440  
 Db 4381 TTCTATGACCAAAACCGTGGGCTTCAAGTCAAGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4440  
 QY 4441 TCAGAGCCAGCTGCTGCTACAGGGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4500  
 Db 4441 TCAGAGCCAGCTGCTGCTACAGGGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4500  
 QY 4501 GGTGCTTAATCAGGTGACAGCAGTGGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4560  
 Db 4501 GGTGCTTAATCAGGTGACAGCAGTGGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4560  
 QY 4561 CAGATTGACTTGCATGCCATCAGAGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4620  
 Db 4561 CAGATTGACTTGCATGCCATCAGAGGAGCGGCGGCGGCTGCAAGCTTGGCAGTTGTCAAGGGC 4620  
 QY 4621 CTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGCTCACCACGCGCTCGG 4680  
 Db 4621 CTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGCTCACCACGCGCTCGG 4680

QY 4681 GCGTCCCTCCCATCCAGTCCGCTGCTCCATGAGCACCACCAACATGGCTATCGGGAGCAT 4740  
 Db 4681 GCGTCCCTCCCATCCAGTCCGCTGCTCCATGAGCACCACCAACATGGCTATCGGGAGCAT 4740  
 QY 4741 GAGTTCTTTGCTGAGCTCCCTAGCGGAAAGAAAGCAAAATTCCTTGCAATAGGC 4800  
 Db 4741 GAGTTCTTTGCTGAGCTCCCTAGCGGAAAGAAAGCAAAATTCCTTGCAATAGGC 4800  
 QY 4801 TTAGGAAAAAAGAAAGCTGCAACCAACGGAATCAATAGGAGTTGAAGAGATTAACTGAC 4860  
 Db 4801 TTAGGAAAAAAGAAAGCTGCAACCAACGGAATCAATAGGAGTTGAAGAGATTAACTGAC 4860  
 QY 4861 TTTGTTTGGCTGTTTTTTAGTTCTGTATGTTTTAGCAATCTCACTCACTCACTCACTCACT 4920  
 Db 4861 TTTGTTTGGCTGTTTTTTAGTTCTGTATGTTTTAGCAATCTCACTCACTCACTCACTCACT 4920  
 QY 4921 AGATGTTTCAATATATTTCTTTTATGGAAGAGGACTCTGAAAAACCCCTAAAGTATTC 4980  
 Db 4921 AGATGTTTCAATATATTTCTTTTATGGAAGAGGACTCTGAAAAACCCCTAAAGTATTC 4980  
 QY 4981 TAGGAGAAAGCTGCTTCCATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 5040  
 Db 4981 TAGGAGAAAGCTGCTTCCATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 5040  
 QY 5041 CTTTTTAAAAA 5055  
 Db 5041 CTTTTTAAAAA 5055  
 RESULT 2  
 AAL45547  
 ID AAL45547 standard; cDNA; 5055 BP.  
 AC AAL45547;  
 XX  
 XX 11-JUN-2002 (first entry)  
 DT  
 DE Human Gli3 coding sequence SEQ ID NO: 22.  
 XX  
 KW Gli3; screening method; bone induction; cartilage induction;  
 KW orthopaedic disease; dental disease; osteoporosis; hyperosteoarthritis;  
 KW osteopathic; ankylosing spondylitis; osteoporosis; hyperosteoarthritis;  
 KW hyperosteoarthritis; osteoporosis; hyperosteoarthritis; osteoporosis;  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 55..4845  
 FT /\*tag= a /product= "Gli3"  
 FT  
 XX WO200211752-A1.  
 PN  
 XX 14-FEB-2002.  
 PD  
 XX 03-AUG-2001; 2001WO-JP06688.  
 PF  
 XX 04-AUG-2000; 2000JP-0242767.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Hikichi Y;  
 PI  
 XX WPI; 2002-241709/29.  
 DR P-PSDB; AAO17114.  
 DR  
 XX Promotion of bone and cartilage formation using Gli3 protein or DNA  
 PT encoding it for treatment of skeletal disorders  
 XX  
 XX Disclosure; Page 144-147; 154pp; Japanese.  
 PS  
 XX The present invention relates to agents for the promotion of bone and  
 CC cartilage formation which contain as the active component a Gli3 protein

CC or a DNA encoding a Gli1 protein. The agents can be used in the  
 CC prevention, treatment and diagnosis of bone and cartilage disorders  
 CC including bone fractures, joint deformation, osteoarthritis,  
 CC osteoporosis, cartilage damage, trauma, bone formation defects, cartilage  
 CC formation defects, bone defects, dental disease, hyperosteoarthritis and  
 CC hyperchondrogenesis, and for use in cosmetic and therapeutic bone  
 CC transplantation. The present sequence is a human Gli3 coding sequence  
 CC described in the exemplification of the invention.

XX Sequence 5055 BP; 1243 A; 1591 C; 1290 G; 931 T; 0 other;

SQ Query Match 100.0%; Score 5055; DB 24; Length 5055;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTACTAGTGGGCAATTTTGGTCAAGAGAGCTGAAGTAATGAGACATCATGAG 60  
 DB 1 CGTACTAGTGGGCAATTTTGGTCAAGAGAGCTGAAGTAATGAGACATCATGAG 60  
 QY 61 GCCAGTCCACAGCTCCACGACCTGAAAGAAAAAGTTGAGAAATCCATAGTGAAG 120  
 DB 61 GCCAGTCCACAGCTCCACGACCTGAAAGAAAAAGTTGAGAAATCCATAGTGAAG 120  
 QY 121 TGCTCCACTCGACAGATGTGCGGAGAAAGCGTTGCTCCAGCACCACCTTCTAATGAG 180  
 DB 121 TGCTCCACTCGACAGATGTGCGGAGAAAGCGTTGCTCCAGCACCACCTTCTAATGAG 180  
 QY 181 GATGAAGTCTCGGAGAGATTTATCACAGAGAGAGAAAGCAATCACTATGAGGCA 240  
 DB 181 GATGAAGTCTCGGAGAGATTTATCACAGAGAGAGAAAGCAATCACTATGAGGCA 240  
 QY 241 CAGAATGTCCAGGGCTCAGCAAGTCACTGAGGACCTTCAACATCGAGTCAAGAGG 300  
 DB 241 CAGAATGTCCAGGGCTCAGCAAGTCACTGAGGACCTTCAACATCGAGTCAAGAGG 300  
 QY 301 GCCTCATTTGATCAAGAAAGATGCCATGGTCCCTGCCACAGTGGCGAGCGCTCTGTG 360  
 DB 301 GCCTCATTTGATCAAGAAAGATGCCATGGTCCCTGCCACAGTGGCGAGCGCTCTGTG 360  
 QY 361 CGGTACCGGGGAGCGTGTGGTTCATGAGACCCAGGAATGGTTACATGAGGCGCCACTAC 420  
 DB 361 CGGTACCGGGGAGCGTGTGGTTCATGAGACCCAGGAATGGTTACATGAGGCGCCACTAC 420  
 QY 421 CACCTCTCTCATCTTTCCCTGCTTCCATCTCCCTGCTGACCAATTCATGCCAGACATCAT 480  
 DB 421 CACCTCTCTCATCTTTCCCTGCTTCCATCTCCCTGCTGACCAATTCATGCCAGACATCAT 480  
 QY 481 GAGGGCGGTACCATTAACGATCCATCTCCGATTCCTCCATTTGCATATGACTTCCGCTTA 540  
 DB 481 GAGGGCGGTACCATTAACGATCCATCTCCGATTCCTCCATTTGCATATGACTTCCGCTTA 540  
 QY 541 TCTAGTAGCCCTACGTATCGGACCTGCGCTTCAATTTAGGATCTCCGCCAACCCGCC 600  
 DB 541 TCTAGTAGCCCTACGTATCGGACCTGCGCTTCAATTTAGGATCTCCGCCAACCCGCC 600  
 QY 601 GCTGCTGCTTCCAGTCTCCCTCAGCCCTCCACATCCCTACATTAATCCCTACATGGAC 660  
 DB 601 GCTGCTGCTTCCAGTCTCCCTCAGCCCTCCACATCCCTACATTAATCCCTACATGGAC 660  
 QY 661 TATATCCGCTCTCTGACAGAGCCCATCGCTCCATGATCTCAACAACCCGCTGGGCTG 720  
 DB 661 TATATCCGCTCTCTGACAGAGCCCATCGCTCCATGATCTCAACAACCCGCTGGGCTG 720  
 QY 721 AGCCCTACAGATGGCCCGCATGAGAGTCCAGGAGTCCAGAGAGATTAATCATAGATGGCC 780  
 DB 721 AGCCCTACAGATGGCCCGCATGAGAGTCCAGGAGTCCAGAGAGATTAATCATAGATGGCC 780  
 QY 781 CTGCTAACTGGCCAGCGCCAGCCCTATGAGAGATTAATCCCTACCTGCCACCCCGGC 840  
 DB 781 CTGCTAACTGGCCAGCGCCAGCCCTATGAGAGATTAATCCCTACCTGCCACCCCGGC 840  
 QY 841 ACGGGGGCCATCCACATGGAATATCTTCATGCTATGATGAGACCATCATCTCCAGCCCC 900  
 DB 841 ACGGGGGCCATCCACATGGAATATCTTCATGCTATGATGAGACCATCATCTCCAGCCCC 900

DB 841 ACGGGGGCCATCCACATGGAATATCTTCATGCTATGATGAGACCATCATCTCCAGCCCC 900  
 QY 901 AGGCTGTACAGCCAGCCGAGCGGAAAGGTACACTGTCATATCACCACTCTCCGATCAT 960  
 DB 901 AGGCTGTACAGCCAGCCGAGCGGAAAGGTACACTGTCATATCACCACTCTCCGATCAT 960  
 QY 961 AGCTTTGACCTTCAGACCATGATAAGGACAGCTCTCCCAACTCCTTTGGTCAAGATTTCAAT 1020  
 DB 961 AGCTTTGACCTTCAGACCATGATAAGGACAGCTCTCCCAACTCCTTTGGTCAAGATTTCAAT 1020  
 QY 1021 AATTCCTAGCAGCTCTTCAGCAAGTGGCTCCTATGCTCACTTATCTGCAAGTGCATC 1080  
 DB 1021 AATTCCTAGCAGCTCTTCAGCAAGTGGCTCCTATGCTCACTTATCTGCAAGTGCATC 1080  
 QY 1081 AGCCCTGCTTGAAGTTCACCTACTCTTCCGCGCCGCTCTCTCCACATGATCATCAGAG 1140  
 DB 1081 AGCCCTGCTTGAAGTTCACCTACTCTTCCGCGCCGCTCTCTCCACATGATCATCAGAG 1140  
 QY 1141 ATCCTAAGCCGACACAGAGCTTAGGTTAGCTTTCAGCCTTTGGACACAGCCCTCCACTCATCCAC 1200  
 DB 1141 ATCCTAAGCCGACACAGAGCTTAGGTTAGCTTTCAGCCTTTGGACACAGCCCTCCACTCATCCAC 1200  
 QY 1201 CCTGCCCAACTTTTCCAAACACAGAGGCTATTCAGGGATTCCTACGGTTCTGAACCC 1260  
 DB 1201 CCTGCCCAACTTTTCCAAACACAGAGGCTATTCAGGGATTCCTACGGTTCTGAACCC 1260  
 QY 1261 GTCCAGTTCAGCTCGGCGCTTCTGAGTCTTCACAGAAAGCCACAGGCTGAGTCTGCA 1320  
 DB 1261 GTCCAGTTCAGCTCGGCGCTTCTGAGTCTTCACAGAAAGCCACAGGCTGAGTCTGCA 1320  
 QY 1321 GTGACAGCACTGTGTGACCCGATGCACAAAGAGGTCCAAAGATCAAAACCCGATGAAGAC 1380  
 DB 1321 GTGACAGCACTGTGTGACCCGATGCACAAAGAGGTCCAAAGATCAAAACCCGATGAAGAC 1380  
 QY 1381 CTCCCAAGCCAGGGGCTCGGGGCGCAGCAGGAGCAAGCCGAGGAAACACCTTTGTCAAG 1440  
 DB 1381 CTCCCAAGCCAGGGGCTCGGGGCGCAGCAGGAGCAAGCCGAGGAAACACCTTTGTCAAG 1440  
 QY 1441 GAGGAAGGGGCAAAAGATGAAGCAACAGAGGCTGAAGTCACTATGAGACAACTGC 1500  
 DB 1441 GAGGAAGGGGCAAAAGATGAAGCAACAGAGGCTGAAGTCACTATGAGACAACTGC 1500  
 QY 1501 CACTGGGAAGGCTCGGCGAGGAGTTCGACACCAAGAGGAGCTTGTGACCATATAAT 1560  
 DB 1501 CACTGGGAAGGCTCGGCGAGGAGTTCGACACCAAGAGGAGCTTGTGACCATATAAT 1560  
 QY 1561 AACGACCATATTCATGAGAGAGAGGAGTTCGCTGTCAGGCTGGCTGACTGCTCAAGA 1620  
 DB 1561 AACGACCATATTCATGAGAGAGAGGAGTTCGCTGTCAGGCTGGCTGACTGCTCAAGA 1620  
 QY 1621 GAGCAAGAACCTTCAAAGCCAGTATATGTTGTTAGTGCATATGAGAGACACACGGGC 1680  
 DB 1621 GAGCAAGAACCTTCAAAGCCAGTATATGTTGTTAGTGCATATGAGAGACACACGGGC 1680  
 QY 1681 GAGAGACCTCAAAATGCACCTTTTGAAGGTTGCACAAAGGCTTACTCGAGACTGAAAC 1740  
 DB 1681 GAGAGACCTCAAAATGCACCTTTTGAAGGTTGCACAAAGGCTTACTCGAGACTGAAAC 1740  
 QY 1741 TTGAAACACACTTGGAGATCTCACACTGGAGAGAACCATACGTTGTGAGCAGCAAGGT 1800  
 DB 1741 TTGAAACACACTTGGAGATCTCACACTGGAGAGAACCATACGTTGTGAGCAGCAAGGT 1800  
 QY 1801 TGCACAAAGGCTTCTCAATGCTCTGATCGGCGCAACACCAACAGAACGATTC 1860  
 DB 1801 TGCACAAAGGCTTCTCAATGCTCTGATCGGCGCAACACCAACAGAACGATTC 1860  
 QY 1861 AATGAGAACCATATGTTGCAAAATCCAGCTGCATTAAGCGTTTACAGAGCCCAAGC 1920  
 DB 1861 AATGAGAACCATATGTTGCAAAATCCAGCTGCATTAAGCGTTTACAGAGCCCAAGC 1920  
 QY 1921 TCCCTCCGGAACATGTGAAGACAGTGTGATGCCAGAGGCTCATGTCCACCAAGAGCAG 1980  
 DB 1921 TCCCTCCGGAACATGTGAAGACAGTGTGATGCCAGAGGCTCATGTCCACCAAGAGCAG 1980

QY 1981 CGAGGGACATCCATCTCTGGCGCCACCCCGAGAGATTCCGGCAGCCATTTCAGATCC 2040  
Db 1981 CGAGGGACATCCATCTCTGGCGCCACCCCGAGAGATTCCGGCAGCCATTTCAGATCC 2040  
QY 2041 AGGTCCCTGGCCGACCGACTCAGGAGGCCCTTGGTGGAGCAGCAGGACCTCAGCAACT 2100  
Db 2041 AGGTCCCTGGCCGACCGACTCAGGAGGCCCTTGGTGGAGCAGCAGGACCTCAGCAACT 2100  
QY 2101 ACCTCAAGCGGGAAGATTCCTCCAGGTGAAACCGTCAAGGCAGAGAAGCCATGACA 2160  
Db 2101 ACCTCAAGCGGGAAGATTCCTCCAGGTGAAACCGTCAAGGCAGAGAAGCCATGACA 2160  
QY 2161 TCTCAGCAAGCCCTGGTGGTCACTTTCATGACGACGCCAACAGTCCCTCAGT 2220  
Db 2161 TCTCAGCAAGCCCTGGTGGTCACTTTCATGACGACGCCAACAGTCCCTCAGT 2220  
QY 2221 TATTCACAGATGGGCTCGAGCTTCTCTGACCGATGAGGTAGTATGAGGACCTCAGT 2280  
Db 2221 TATTCACAGATGGGCTCGAGCTTCTCTGACCGATGAGGTAGTATGAGGACCTCAGT 2280  
QY 2281 GCCATCGATGAACCCCAATATGACTCAACCATTTCCACTGCAACACACAGCCCTTGCT 2340  
Db 2281 GCCATCGATGAACCCCAATATGACTCAACCATTTCCACTGCAACACACAGCCCTTGCT 2340  
QY 2341 TTGCAAGCCAGGAAACCCCGCAGGAGCCAAATGGATGGAGCAGTAAAGAGG 2400  
Db 2341 TTGCAAGCCAGGAAACCCCGCAGGAGCCAAATGGATGGAGCAGTAAAGAGG 2400  
QY 2401 CTAAACAGTGAATGGATGTTTCGGGACTGAACCCCTTACCCCTTAAGCCCT 2460  
Db 2401 CTAAACAGTGAATGGATGTTTCGGGACTGAACCCCTTACCCCTTAAGCCCT 2460  
QY 2461 GCGGTCTCTCTCATAGGAATGGACACAGTCCCAACACACCTGCAGCTGGGTGG 2520  
Db 2461 GCGGTCTCTCTCATAGGAATGGACACAGTCCCAACACACCTGCAGCTGGGTGG 2520  
QY 2521 CCCATGAGCTTCTCCCGGCGAGAGCAGCTCTCTGGGTGGAGCTGACATGCTGAAC 2580  
Db 2521 CCCATGAGCTTCTCCCGGCGAGAGCAGCTCTCTGGGTGGAGCTGACATGCTGAAC 2580  
QY 2581 ATGCTCAACAGAGGACAGAGCCAGCAGCAGCAGCTGAGCTGAGCTGAGCTGAGCAGCCG 2640  
Db 2581 ATGCTCAACAGAGGACAGAGCCAGCAGCAGCAGCTGAGCTGAGCTGAGCAGCCG 2640  
QY 2641 CGCTCCTCAGGATCTCGCCTGCTTCTCCAGCCGCGCTCCAGGAGGCGTCAAGGCC 2700  
Db 2641 CGCTCCTCAGGATCTCGCCTGCTTCTCCAGCCGCGCTCCAGGAGGCGTCAAGGCC 2700  
QY 2701 GAGGCGCGCCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2760  
Db 2701 GAGGCGCGCCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2760  
QY 2761 TCGCGCGCTCCAGGAGGACGAGCAGGAGCAGGAGCAGCTGAGCTGAGCTGAGCTGAGCTGAG 2820  
Db 2761 TCGCGCGCTCCAGGAGGACGAGCAGGAGCAGGAGCAGCTGAGCTGAGCTGAGCTGAGCTGAG 2820  
QY 2821 CCGCGCCGAGCTACCGCTCAAGGCCAAGTACGGGCTGCCAGAGAGGCGCCGCGCG 2880  
Db 2821 CCGCGCCGAGCTACCGCTCAAGGCCAAGTACGGGCTGCCAGAGAGGCGCCGCGCG 2880  
QY 2881 AGCCCTCCGCAACATGGAGGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2940  
Db 2881 AGCCCTCCGCAACATGGAGGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2940  
QY 2941 GCGCTCGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 3000  
Db 2941 GCGCTCGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 3000  
QY 3001 GGAGCCCGGCTAGCGGCGGCGCCACCTGACGCGGACGATGCGCTGGGCGAGCGGCTG 3060  
Db 3001 GGAGCCCGGCTAGCGGCGGCGCCACCTGACGCGGACGATGCGCTGGGCGAGCGGCTG 3060

QY 3061 AGGAGGCCAGCGACCGGTCGGGACAGGTCGAGAGGCTGCGCCCTGCTCGTGTGCGG 3120  
Db 3061 AGGAGGCCAGCGACCGGTCGGGACAGGTCGAGAGGCTGCGCCCTGCTCGTGTGCGG 3120  
QY 3121 CGCTTCAGCAGCCTCAGCAGCTGCAACCCCGCGGATGCGCAGCTCCGCGGAGAGCGC 3180  
Db 3121 CGCTTCAGCAGCCTCAGCAGCTGCAACCCCGCGGATGCGCAGCTCCGCGGAGAGCGC 3180  
QY 3181 AGTCTCGTGTTCAGAAATTACAGCGCGCGAGGCGGCGCAGTCCGGAATTTCCACTCG 3240  
Db 3181 AGTCTCGTGTTCAGAAATTACAGCGCGCGAGGCGGCGCAGTCCGGAATTTCCACTCG 3240  
QY 3241 TCCCTCTCTCTCCAGCATCAGGAGAGCTCAACCTGGAGTCCCTGACCATGGAGCT 3300  
Db 3241 TCCCTCTCTCTCCAGCATCAGGAGAGCTCAACCTGGAGTCCCTGACCATGGAGCT 3300  
QY 3301 GATGCCAACCTGAACGATGAGGATTTCTCGCGGACGACGTGGTGCAGTATTTAAATTC 3360  
Db 3301 GATGCCAACCTGAACGATGAGGATTTCTCGCGGACGACGTGGTGCAGTATTTAAATTC 3360  
QY 3361 CAGAACCAAGCAGGATACGAGCAGCAGCTTCCCGAGCGCCCTCCCGGACGACACCAAGTG 3420  
Db 3361 CAGAACCAAGCAGGATACGAGCAGCAGCTTCCCGAGCGCCCTCCCGGACGACACCAAGTG 3420  
QY 3421 CCCACGGGCGCGGTGACTTTGACGCGCCCGGCTGCGCAGACAGCCAGCTGCCAGCAG 3480  
Db 3421 CCCACGGGCGCGGTGACTTTGACGCGCCCGGCTGCGCAGACAGCCAGCTGCCAGCAG 3480  
QY 3481 TTCATGCCCTCAGCAGCAGCTTCCCGGAGGCGAGCAAAACCGACCTGCCATTCAGTGG 3540  
Db 3481 TTCATGCCCTCAGCAGCAGCTTCCCGGAGGCGAGCAAAACCGACCTGCCATTCAGTGG 3540  
QY 3541 AACGAAGTACGCTCCGGAAGCGCGACTGCTCTCTCCAAAGCTCAAGTGTGGCGCGCG 3600  
Db 3541 AACGAAGTACGCTCCGGAAGCGCGACTGCTCTCTCCAAAGCTCAAGTGTGGCGCGCG 3600  
QY 3601 CCGCTGTGCGGAGCTCGCGCTTTGGGTTCTGAAAGGATGCTGCTCCAGCTGAGTGTGGCGCG 3660  
Db 3601 CCGCTGTGCGGAGCTCGCGCTTTGGGTTCTGAAAGGATGCTGCTCCAGCTGAGTGTGGCGCG 3660  
QY 3661 AACCTTCAGGAGCGGCTGCTGGGGCTATCAGACCTCGGGGAGACACACACCCC 3720  
Db 3661 AACCTTCAGGAGCGGCTGCTGGGGCTATCAGACCTCGGGGAGACACACACCCC 3720  
QY 3721 TAGGTTGGCCAGAGCAGCTTATGCTCCCAACAGCCCGGAAAGTGCACAGTGGAAAC 3780  
Db 3721 TAGGTTGGCCAGAGCAGCTTATGCTCCCAACAGCCCGGAAAGTGCACAGTGGAAAC 3780  
QY 3781 GCCTTCATGAACAGCCTGTAAAGGCGCGCAGTATGGAACCTGCTCAACAGGAGCAG 3840  
Db 3781 GCCTTCATGAACAGCCTGTAAAGGCGCGCAGTATGGAACCTGCTCAACAGGAGCAG 3840  
QY 3841 GTGGCCCTGCTGACTCGAGCTGCTGCTGGTGGCGGATTCAGGCTCAAGCTGAAG 3900  
Db 3841 GTGGCCCTGCTGACTCGAGCTGCTGCTGGTGGCGGATTCAGGCTCAAGCTGAAG 3900  
QY 3901 AGACCCCATCAAGGAGCGGCGCAGCTGAATTTTCGGCTGCGGTAGCGCCCAAT 3960  
Db 3901 AGACCCCATCAAGGAGCGGCGCAGCTGAATTTTCGGCTGCGGTAGCGCCCAAT 3960  
QY 3961 GAGTACGCTGGCAGCATGTTGATGCGATGCAAGAACAGGACCTGAGGAGGAGGTAC 4020  
Db 3961 GAGTACGCTGGCAGCATGTTGATGCGATGCAAGAACAGGACCTGAGGAGGAGGTAC 4020  
QY 4021 CTGGCTCAGCAGCTCCTCGCGACAGCATGCGACCCCGGGGCGAGCGCCCGGCTCAG 4080  
Db 4021 CTGGCTCAGCAGCTCCTCGCGACAGCATGCGACCCCGGGGCGAGCGCCCGGCTCAG 4080  
QY 4081 CAGATGCTGGGCGAGATTAGTGTACTCTACATCAACATCTTACCAAGGCGCAGAGAGC 4140  
Db 4081 CAGATGCTGGGCGAGATTAGTGTACTCTACATCAACATCTTACCAAGGCGCAGAGAGC 4140  
QY 4141 TGCTGCGGCGGCTCAGGCGATGGGCGAGCCGCTGCAAGCTTGGCAGTTGTGAGGGGC 4200

```

Db 4141 TGCTGCCAGGGGCTCAGGCGATGGCAGCCAGCCGCTCAAGCTGGCAGTTGTCCAGGGC 4200
Qy 4201 TACAGCCCATGTGCAGCTTTGGGGGCGAGCAGGGCCAGGCTATGCCAGGACAGCCCTT 4260
Db 4201 TACAGCCCATGTGCAGCTTTGGGGGCGAGCAGGGCCAGGCTATGCCAGGACAGCCCTT 4260
Qy 4261 GCTCTGCAGTCAGGACAGCTCAGTCAGCAAGTCAGAGCTGCAGGTCGAATGGTATCAAG 4320
Db 4261 GCTCTGCAGTCAGGACAGCTCAGTCAGCAAGTCAGAGCTGCAGGTCGAATGGTATCAAG 4320
Qy 4321 ATGGAGATGAAGAGGCGAGCCCATCCGCTGTCTCTAATCTGCAGAAATTAATCTCTGTCAG 4380
Db 4321 ATGGAGATGAAGAGGCGAGCCCATCCGCTGTCTCTAATCTGCAGAAATTAATCTCTGTCAG 4380
Qy 4381 TTCTATGACCAACCGTGGGCTTCAGTCAGCAAGACAGAAAGCTGGTTCATCTCTATT 4440
Db 4381 TTCTATGACCAACCGTGGGCTTCAGTCAGCAAGACAGAAAGCTGGTTCATCTCTATT 4440
Qy 4441 TCAGAGCCGAGCTGCTGTCTACAGGGGACCGAGCCGCAAAAACCTCTGAGTTACTTTCCCCA 4500
Db 4441 TCAGAGCCGAGCTGCTGTCTACAGGGGACCGAGCCGCAAAAACCTCTGAGTTACTTTCCCCA 4500
Qy 4501 GGTGCTAATCAGGTGACAGCAGCTGGACAGCCCTCGACAGCCATGACCTGGAAGGGGTA 4560
Db 4501 GGTGCTAATCAGGTGACAGCAGCTGGACAGCCCTCGACAGCCATGACCTGGAAGGGGTA 4560
Qy 4561 CAGATTGACTTCGATGCCCATATAGACGATGGGGACCACTCCAGCCTGATGTCGGGGGCC 4620
Db 4561 CAGATTGACTTCGATGCCCATATAGACGATGGGGACCACTCCAGCCTGATGTCGGGGGCC 4620
Qy 4621 CTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGCTCACCAGCCTCGG 4680
Db 4621 CTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGCTCACCAGCCTCGG 4680
Qy 4681 GGTGCTCCTCCCATTCGCTGCTGTCCATGAGCACCACCAACATGGCTATCGGGGACAT 4740
Db 4681 GGTGCTCCTCCCATTCGCTGCTGTCCATGAGCACCACCAACATGGCTATCGGGGACAT 4740
Qy 4741 GAGTTCTTTGCTGACCTCCCTAGCGGAGAAAGCAAAATTCCTTGCAATGCAATAGGC 4800
Db 4741 GAGTTCTTTGCTGACCTCCCTAGCGGAGAAAGCAAAATTCCTTGCAATGCAATAGGC 4800
Qy 4801 TTTAGAAAAAAGACTGCAACCAACGGAATCAATAGGAGTTGAAGATTAAACTGAC 4860
Db 4801 TTTAGAAAAAAGACTGCAACCAACGGAATCAATAGGAGTTGAAGATTAAACTGAC 4860
Qy 4861 TTTGTTTTGGCTGTTTTTTTAGTTCTGTATGTTATTTAGCAATCTCATCTACCTTAAC 4920
Db 4861 TTTGTTTTGGCTGTTTTTTTAGTTCTGTATGTTATTTAGCAATCTCATCTACCTTAAC 4920
Qy 4921 AGATGTTTTCAATATATATTTTATGGAAAAAGGACTCTGAAAAACCTTAAGATTTC 4980
Db 4921 AGATGTTTTCAATATATATTTTATGGAAAAAGGACTCTGAAAAACCTTAAGATTTC 4980
Qy 4981 TAGGGAGAAACTGCTCTCCATTTTCAATTTTGAATCAGTATTTTACACTCAAAACACCT 5040
Db 4981 TAGGGAGAAACTGCTCTCCATTTTCAATTTTGAATCAGTATTTTACACTCAAAACACCT 5040
Qy 5041 CTTTTTAAAAA 5055
Db 5041 CTTTTTAAAAA 5055

RESULT 3
ID AAL45546
ID AAL45546 standard; cDNA; 5113 BP.
XX
AC AAL45546;
XX
XX 11-JUN-2002 (first entry)
XX
DE Murine Gli3 coding sequence SEQ ID NO: 20.

```

```

XX Gli1: screening method; bone induction; cartilage induction;
KW orthopaedic disease; dental disease; osteoporosis; hyperosteoecgenesis;
KW osteopathic; antiarthritic; vulnery; immunosuppressive;
KW hyperchondrogenesis; mouse; gene; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 172..4962
FT /*tag= a
FT /product= "Gli3"
FT /note= "q"
XX
PN W0200211752-A1.
XX
PD 14-FEB-2002.
XX
PF 03-AUG-2001; 2001WO-JP06688.
XX
PR 04-AUG-2000; 2000JP-0242767.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Hikichi Y;
XX
XX WPI: 2002-241709/29.
DR P-PSDB; AAO17113.
XX
XX Promotion of bone and cartilage formation using Gli1 protein or DNA
XX encoding it for treatment of skeletal disorders
XX
XX Disclosure; Page 133-136; 154pp; Japanese.
XX
XX The present invention relates to agents for the promotion of bone and
XX cartilage formation which contain as the active component a Gli1 protein
XX or a DNA encoding a Gli1 protein. The agents can be used in the
XX prevention, treatment and diagnosis of bone and cartilage disorders
XX including bone fractures, joint deformation, osteoarthritis,
XX osteoporosis, cartilage damage, trauma, bone formation defects, cartilage
XX formation defects, bone defects, dental disease, hyperosteoecgenesis and
XX hyperchondrogenesis, and for use in cosmetic and therapeutic bone
XX transplantation. The present sequence is a murine Gli3 coding sequence
XX described in the exemplification of the invention.
XX
XX Sequence 5113 BP; 1337 A; 1490 C; 1254 G; 1032 T; 0 other;
XX
Query Match 73.1%; Score 3696.6; DB 24; Length 5113;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 4212; Conservative 0; Mismatches 769; Indels 24; Gaps 3;
Qy 9 CTGGGCAATTTTGGTCCGAGAGAGCTGAAGTAATGAGAGACATCATGAGGCCAGTC 68
Db 126 CGCAGGGATTCTCTTTGAGAAAACAAGCTGAAGTAATGAGAGACATTTATGAGGCCAGGC 185
Qy 69 CCACAGCTCCACGACCACTGAAAGAAAAGTTGAGAAATTCATAGTGAAGTCTCCAC 128
Db 186 CCACAGCTCTACGGGACTGAGAGAGAGAAAGCTGAAATTCATTTGGAAATGCTCCAC 245
Qy 129 TCGAACAGATGTGAGCGAGAGAAAGCGCTTGCCTCCAGCACCACCTTCTTAATGAGGATGAAG 188
Db 246 GAGAACAGATGTGAGCGAGAGAGCGCTTGCCTCCAGCACCACCTTCTTAATGAGGATGAAG 305
Qy 189 TCCTGGACAGACTTATCAGAGAGAGAGAAAGCAATCACTATGAGCCACAGATGT 248
Db 306 TCCTGGACAGACTTATCAGAGAGAGAGAAAGCAATCACTATGAGCCCTCAGAGT 365
Qy 249 CCAGGGCTCAGCAAGATCAGTGAAGAACCTTCAACATCGAGTCAGAGAGGCGCTCAT 308
Db 366 GCAGGGTCTCAACAAATCAGTGAAGAGCCCTCGACGCTCTAGTATGAGAGGCGCTCGCT 425
Qy 309 GATCAAGAAAGAGATCCATGGTCCCTGCCACAGCTGCGGAGCCCTCTGTGCGGTACCG 368

```







Db 4806 AFCCCTCCATCCCAATCCCTATCCATGGGACCAACACATGGCTATCGGGATATGA 4865  
 Qy 4743 GTTCCTTCTGCTGACCTCCCTAGCGAAGAACAAATTCCTTGCAGTTATGCAATAGGCTT 4802  
 Db 4866 GTTCCTTCTGCTGACCTCCCTTGCAGAAAGAACAAATTCCTTGCAGTTATGCAATAGGCTT 4925  
 Qy 4803 TAGGAAAAAAGACTGCAACCAACGGAATCAATAGGAGTTGAAGAGATTAAACTGACTT 4862  
 Db 4926 TAGGCAAGAGGACCAACAAACAAAGACTGAATGACTTGGGA----- 4969  
 Qy 4863 TGTCTTGGCTGTTTTTTAGTTCTGTATGTAATTTAGCAATCTCATCTCACTCACTGAG 4922  
 Db 4970 -TTTTTTTTCTCTTTTAAAGTCTGTGTATTTTAGCAATCTCATCTCACTCACTGAGG 5028  
 Qy 4923 ATGTCTTCAATATATTCCTTTTATGAAAGGACTCTGAAACCCCTAAAGTATTTCTA 4982  
 Db 5029 ATGTCTCTCAAGTATATTCCTTTTATGAAAGGACTCTGAAACCCCTAAAGTATTTCTA 5088  
 Qy 4983 -GGGAGAAACTGCTTCCATTTTCAG 5006  
 Db 5089 GGGGAGAAACTGCTTCCATTTTCAG 5113

RESULT 4  
 ABQ15580/C  
 ID ABQ15580 standard; DNA; 557 BP.  
 XX  
 AC ABQ15580;  
 XX  
 12-JUL-2002 (first entry)  
 XX  
 Oligonucleotide for detecting cytosine methylation SEQ ID NO 2171.  
 XX  
 Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 drug; side effect; cancer; central nervous system; cardiovascular;  
 gastrointestinal; respiratory system; single nucleotide polymorphism;  
 SNP; cell differentiation; ds.  
 XX  
 Homo sapiens.  
 XX  
 WO200218632-A2.  
 XX  
 07-MAR-2002.  
 XX  
 01-SEP-2001; 2001WO-EP10074.  
 XX  
 01-SEP-2000; 2000DE-1043826.  
 XX  
 05-SEP-2000; 2000DE-1044543.  
 XX  
 (EPIG-) EPIGENOMICS AG.  
 XX

Olek A, Piepenbrock C, Berlin K, Guetig D;  
 WPI; 2002-371829/40.

Determining the degree of cytosine methylation in genomic DNA, useful  
 for diagnosis and prognosis, comprises selective hybridization of  
 amplicons from chemically treated DNA

Claim 12; 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of  
 methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 genomic sample of DNA. The sample is treated chemically to convert  
 cytosine (C) but not methylated C, to uracil, then part of the genomic  
 DNA that contains the target C is amplified to form a labeled amplicon.  
 The amplicon is hybridised to two classes, each with at least one  
 member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 and the degree of hybridisation to both classes is determined from the  
 label on the amplicon. From the ratio of labels hybridised to the two  
 classes of oligomers, the degree of methylation is calculated. The method  
 is used: (i) for diagnosis and/or prognosis of side effects of  
 therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders

CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in  
 CC the disclosure of the invention.

XX Sequence 557 BP; 75 A; 52 C; 205 G; 225 T; 0 other;

Query Match 7.2%; Score 366; DB 24; Length 557;  
 Best Local Similarity 79.1%; Pred. No. 2.4e-72;  
 Matches 435; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy 2989 TGACGACGGGGGAGCCACGCTAGCGGCGGCGCCACTGCGAGCGCAGGATGGCTG 3048  
 Db 552 TACAACGACGAAAAAAGCCACGACTAGCAGACGACCCACTACACCCGACGATACGCG 493  
 Qy 3049 GGCACGCGGTGAGGAGGCGCAGCGCCGCTCGGACAGGCTCCGAGGGGCTGGGCCCTG 3108  
 Db 492 AACACGACGTAAAAAACCACGACCCGATACGACAACTCCGAAAACTTAACCTTA 433

Qy 3109 CCTCGTGTGCGCGCTTACGACGCTTACGAGCTCAACACCCCGCGGCGATGCGCAGTCC 3168  
 Db 432 CCTCGTATACCGCGCTTCAACAACTCAACAACTAGAACCCCGCGACGATACGAGTCC 373

Qy 3169 GCGGAGAACGCGAGTCTCGTCTCAGAAATACAGCGGCGCGGCGGCGGCGGCGGCGG 3228  
 Db 372 GCGAAAAACGCAATCTCGTACTTCAAAATACACGCGCCCGGAAAAACGACCAATCCC 313

Qy 3229 AACTTCCACATCTGCTCCCTGCTCTCCAGCATCAGCGAAGACGTCACCTTGGAGTCCCTG 3288  
 Db 312 AACTTCCACATCTGCTCCCTGCTCTCCAGCATCAGCGAAGACGTCACCTTGGAGTCCCTG 253

Qy 3289 ACCATGACGCTGATGCGCAACCTGAAAGATGAGGATTTCTGCGGACGACGCTGGTGAG 3348  
 Db 252 ACCATAAACGCTAATACCAACCTAAACGATAAAAATTTCTTACCGAACGACGCTAAATAC 193

Qy 3349 TATTTAAATTTCCAGAACCAAGGCTACGAGCAGCTTCCCGCGGCGCTCCCGGAC 3408  
 Db 192 TATTTAAATTTCCAGAACCAAGGCTACGAGCAGCTTCCCGCGGCGCTCCCGGAC 133

Qy 3409 GACAGCAAGTGGCCCGCGGCGCTGACTTTGACGCGCGCGGCTGCGACAGACGCCAC 3468  
 Db 132 GACAAACAAATACCCACGAAACCCGATACTTTAAACGCGCGCGGCTTACCAAAACACCA 73

Qy 3469 GCTGGCCGACGATTCATGCGCTTCGAGCAGCGCTGCGGCGGCGGCGGCGGCGGCGG 3528  
 Db 72 GCTAACCAACAATTCATACCTCGAACACCTTACCCCGGAAAAACAAACAAACCGACCTA 13

Qy 3529 CCCATTCACT 3538  
 Db 12 CCCATTCAAT 3

RESULT 5

ABQ15581

ID ABQ15581 standard; DNA; 557 BP.

XX ABQ15581;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



PN WO200218632-A2.  
 XX 07-MAR-2002.  
 XX 01-SEP-2001; 2001WO-EP10074.  
 XX 01-SEP-2000; 2000DE-1043826.  
 PR 05-SEP-2000; 2000DE-1044543.  
 XX (EPIG-) EPIGENOMICS AG.  
 XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
 XX WPI; 2002-371829/40.  
 XX Determining the degree of cytosine methylation in genomic DNA, useful  
 PT for diagnosis and prognosis, comprises selective hybridization of  
 PT amplicons from chemically treated DNA -  
 XX  
 XX Claim 12; 56pp + Sequence Listing; 56pp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one  
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridisation to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridised to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used; (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in  
 CC the disclosure of the invention.  
 XX  
 XX Sequence 557 BP; 225 A; 205 C; 52 G; 75 T; 0 other;  
 SQ  
 Query Match 7.2%; Score 366; DB 24; Length 557;  
 Best Local Similarity 79.1%; Pred. No. 2.4e-72;  
 Matches 435; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
 QY 2989 TGCAGGCGAGGGGAGCCACGCGTACGGCGCGGCCACCTGCAGCGCCGATCGCGTG 3048  
 DB 6 TACACGACGAGAAACCCCGACTACGAAACGACGCCCTACACCGCGCAGATACGCGG 65  
 QY 3049 GGCACGCGGTGAGGAGGGCCAGCGCGTCCGAGACAGGCTCCGAGGGCGTGGCCCTG 3108  
 DB 66 AACGACGAGTAAACAAACACGACGCGGATACGACAACTCCGAAACCTACACCTA 125  
 QY 3109 CTTCTGTGCGCGGCTTACGAGCGCTACGAGCTGCAACCCCGCGGATGCCAGTCC 3168  
 DB 126 CTTCTGTATACCGGCTTCAACACCTCAACAACTCAACACCCCGCGATACACCGTCC 185  
 QY 3169 GCGGAGACGCGAGTCTCTGCTTACAGATTACACCGCGCGCGCGCGGCGCGAGTCCCGA 3228  
 DB 186 GCGAAACACGGAATCTCTGCTTCAAAATACAGCGACCGGAAACGACCAATCTCCGA 245  
 QY 3229 AACTTCCACTCGTCCCGCTCTCCGACATCACCAGAACGTCACCTGGAGTCCCTG 3288  
 DB 246 AACTTCCACTCGTCCCGCTCTCCGACATCACCAGAACGTCACCTGGAGTCCCTG 305  
 QY 3289 ACCATGAGCGTGTGCGGCTGACGAGTACGAGGATTCTCTCGCGGACGAGTGTGTGAG 3348  
 DB 306 ACCATAACGCTAATACCAACCTAAACGATAAAATTTCTACCGAAGCAGCTAATACAA 365  
 QY 3349 TATTTAAATCCAGNACCAACGAGGATACGAGCAGCTTCCCGACGCCCTCCCGAC 3408

Db 366 TATTTAAATCCCAAAACCAAAATACGAAACACTTCCCAACGCCCTCCCGAAC 425  
 QY 3409 GACAGCAAAAGTCCCGACGCGCGGTGACTTTGACGCGCGCGCGGTGCGCAGACGAC 3468  
 Db 426 GACAACAAATACCCGACGAACTTTAAACGCGCGGAACTTACCAAAACGAC 485  
 QY 3469 GCTGGCCAGCAGTTCATGCGCTCGAGCGCCCTGCGCGCGCGCGGAGGAGCAAAACGACCTG 3528  
 Db 486 GCTAACCAAAATTCATACCTCGAACACCTTACCCCGAAACAAACGACCTA 545  
 QY 3529 CCCATTTCAGT 3538  
 Db 546 CCCATTTCAT 555  
 RESULT 6  
 AAL45541  
 ID AAL45541 standard; cDNA; 3635 BP.  
 XX  
 AC AAL45541;  
 XX  
 DT 11-JUN-2002 (first entry)  
 XX  
 DE Murine Gli1 coding sequence SEQ ID NO: 9.  
 XX  
 KW Gli1; screening method; bone induction; cartilage induction;  
 KW orthopaedic disease; dental disease; osteoporosis; hyperosteoecgenesis;  
 KW osteopathic; antiarthritic; vulnary; immunosuppressive; human;  
 KW hyperchondrogenesis; gene; mouse; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 213..3548  
 FT /\*tag= a  
 FT /product= "Gli1"  
 XX  
 PN WO200211752-A1.  
 XX  
 PD 14-FEB-2002.  
 XX  
 XX 03-AUG-2001; 2001WO-JP06688.  
 XX  
 PR 04-AUG-2000; 2000JP-0242767.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Hikichi Y;  
 XX  
 DR WPI; 2002-241709/29.  
 DR P-PSDB; AAO17108.  
 XX  
 PT Promotion of bone and cartilage formation using Gli1 protein or DNA  
 PT encoding it for treatment of skeletal disorders -  
 XX  
 PS Claim 6; Page 88-89; 154pp; Japanese.  
 XX  
 CC The present invention relates to agents for the promotion of bone and  
 CC cartilage formation which contain as the active component a Gli1 protein  
 CC or a DNA encoding a Gli1 protein. The agents can be used in the  
 CC prevention, treatment and diagnosis of bone and cartilage disorders  
 CC including bone fractures, joint deformation, osteoarthritis,  
 CC osteoporosis, cartilage damage, trauma, bone formation defects, cartilage  
 CC formation defects, bone defects, dental disease, hyperosteoecgenesis and  
 CC hyperchondrogenesis, and for use in cosmetic and therapeutic bone  
 CC transplantation. The present sequence is a murine Gli1 coding sequence  
 CC described in the exemplification of the invention.  
 XX  
 SQ Sequence 3635 BP; 786 A; 1140 C; 989 G; 719 T; 1 other;  
 Query Match 6.6%; Score 334.8; DB 24; Length 3635;  
 Best Local Similarity 70.4%; Pred. No. 4.9e-65;











CC tumour formation. The antisense compounds are safely and effectively  
 CC administered to humans. The present sequence represents human  
 CC glioma-associated oncogene-1 DNA.

XX Sequence 3600 BP; 785 A; 1161 C; 949 G; 705 T; 0 other;

Query Match 6.5%; Score 329.4; DB 24; Length 3600;  
 Best Local Similarity 70.6%; Pred. No. 7.9e-64;  
 Matches 454; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

QY 1372 GATGAAGACCTCCCGAGGAGGCTGGGCGACAGACAGACCCGGAAGACAC 1431  
 DB 664 GAAGGTATATGTCCAGCCCACTCCACACAGCATACAGATCCCTGGTGGATCTGT 723  
 QY 1432 CTGTCAAGAGAGAGGAGGACAAAGATGAAGCAAGAGGCTGAGATCATATAGAG 1491  
 DB 724 GATGGGGGAGGAGACCTCGAAGAGAGAGG---GAGCGCTGGCTGATCTGTATGAA 780  
 QY 1492 ACAAACTGCCACTGGGAGGAGGCTGGCGAGGAGTTCGACACCCAGAGAGCTGTGCAC 1551  
 DB 781 ACTGACTGCTGGATGGATGGCTGCAGCAGGATTTGACTCCCAAGAGCAGCTGTGCAC 840  
 QY 1552 CATATAATATACGACCATATTCATGAGAGAGAGAGAGATTCGTGCGAGGTGCTGAC 1611  
 DB 841 CACATCAACAGCGACATCCACAGGAGGAGGAGGATTCGTGCTCCACTGGGGGGGC 900  
 QY 1612 TGCTCAAGAGAGAGAGAAACCTTCAGAGCCAGTATATGTTGAGTGCATATAGAGAGA 1671  
 DB 901 TGCTCCAGGAGAGCTGAGGAGGCTTCAGAGCCAGTACATGCTGGGTTCACATGCGAGA 960  
 QY 1672 CACACGGGAGAGAGGCTCAGCAAAATGACCTTTGAGAGTTCACAAAGGCTTACTCGAGA 1731  
 DB 961 CACACTGGCGAGAGAGCCACACAGTGCAGCTTTGAGAGGTCGCGAGAGTACTACTACGC 1020  
 QY 1732 CTAAAAAACTTGAAGAACACACTTATCTACACTGAGAGAGAAACCATACGTGTGAG 1791  
 DB 1021 CTGCAAAACCTGGAAGAGCGACCTCGGTTCACACAGGAGTGAAGCCATCATGTGTGAG 1800  
 QY 1792 CACGAGAGTTCACAAAGGCTTCTCAAAATGCTCTGATGAGCGGCAACCAACAGAGA 1851  
 DB 1081 CACGAGGCTGCTCAAAAGCTTACAGATTCAGATTCAGAGCCAGGCAACGACCAATCGG 1140  
 QY 1852 ACGCATTCATGAGAAACCATATGTGCAAAATCCAGAGCTTCACTAAGCCTTACAGA 1911  
 DB 1141 ACCCATTCATGAGAGAGCGGTATGTATGATAGCTCCCTGGCTGCACAAACGCTATACA 1200  
 QY 1912 GACCAAGCTCCCTCCGGAACATGTGAGAGAGTGCATGGCCGAGAGGCTCATGTAC 1971  
 DB 1201 GATCCTAGCTGCTGCGAAGAAACATGTCAAGACAGTGCATGTGCTGAGCGCATGTGAC 1260  
 QY 1972 AAGAAGCAGGAGGAGCATCATCTCGGCGCCGACCCCGA 2014  
 DB 1261 AAAGGCAACGCTGGGATGGCCCTCTGCTGGGACACATCA 1303

RESULT 14  
 ABO15578  
 ID ABO15578 standard; DNA: 557 BP.

AC ABO15578;  
 XX 12-JUL-2002 (first entry)

DE oligonucleotide for detecting cytosine methylation SEQ ID NO 2169.

XX Human: cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.

OS Homo sapiens.  
 XX  
 PN W0200218632-A2.

XX 07-MAR-2002.  
 PD  
 XX  
 PF 01-SEP-2001; 2001WO-EP10074.  
 XX  
 PR 01-SEP-2000; 2000DE-1043826.  
 PR 05-SEP-2000; 2000DE-1044543.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guelig D;  
 PI WPI; 2002-371829/40.

PT Determining the degree of cytosine methylation in genomic DNA, useful  
 PT for diagnosis and prognosis, comprises selective hybridization of  
 PT amplicons from chemically treated DNA

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridized to two classes, each with at least one  
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridization to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridized to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used: (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP/s); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC ABO13410-AB054121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in  
 CC the disclosure of the invention.

CC Sequence 557 BP; 105 A; 52 C; 172 G; 228 T; 0 other;

Query Match 6.1%; Score 310.6; DB 24; Length 557;  
 Best Local Similarity 72.4%; Pred. No. 6.7e-60;  
 Matches 403; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 2984 GGAGGTGACGAGAGGAGAGCCACGCTACGAGGCGGCGCCACCTGACCGCAGATG 3043  
 DB 1 GGAGGTGACGAGAGGAGAGGAGTTACGTTACGAGGCGGCTATTGTGTGCTACAGATG 60  
 QY 3044 CGCTGGGCGACGCGGTGAGAGAGGCGGAGCCGCTGCGAGAGGCTCGAGAGGCTG 3103  
 DB 61 CGTGGGTTACGCGGTGAGAGAGGAGGATTCGCTGCGGATGAGGTTGAGAGGTTGG 120  
 QY 3104 CCTGCGCTGCTGGCGGCTTCAGAGCCTCAGACGCTGCAACCCCGGGGATGGGCA 3163  
 DB 121 TTTTGTGTTCTGTGTCGCTTTTAGTAGTTTAGTAGTTTAGTTTTCGGGAGATGTA 180  
 QY 3164 CGTCCGCGAGAGAGCGCAGTCTGCTTCAGAAATTCAGAGGCGCCGAGAGGCGGCA 3223  
 DB 181 CGTTGCGGAGAGAGCGTACGTTGCTTTTAGAATATATACGGGCTTCAGAGGCGGTA 240  
 QY 3224 CCGGAATCTCCACTGCTCCCTGCTGCTCCAGCATACCGAGAGAGCTACCTGAGAT 3283  
 DB 241 TTGGAATTTTATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 300  
 QY 3284 CCTGAGCATGAGCGGTGATGCAACCTGACAGATAGAGATTTCCGCGCGAGAGCGTGG 3343  
 DB 301 TTTTGATTTATGAGCTTTGATTTAATTTAGACAGATGAGATTTTGTGAGAGAGTGG 360  
 QY 3344 TGCAGTATTTAATTCGAGAACAGAGGAGTACGAGAGCAGCTCCCGAGCGCTCC 3403







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2002, 20:45:36 ; Search time 111 seconds  
(without alignments)

13966.226 Million cell updates/sec

Title: US-09-910-185-3

Sequence: 1 cgatactagctggcattt.....accctcttttaaaaaa 5055

Scoring table: IDENTITY\_NIC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

1: Issued Patents\_NA:  
2: /cgn2\_6/ptodata/1/lna/5A.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/lna/5B.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/lna/6A.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/lna/6B.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/lna/PTUTS.COMB.seq:\*  
7: /cgn2\_6/ptodata/1/lna/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	14.0	4960	US-09-907-843-3	Sequence 3, Appl
2	329.4	6.5	3600	US-09-657-042A-3	Sequence 3, Appl
3	90.8	1.8	3138	US-09-234-332-5	Sequence 5, Appl
4	83.2	1.6	2364	US-09-172-045-1	Sequence 1, Appl
5	68.6	1.4	7218	US-08-232-463-14	Sequence 14, Appl
6	61	1.2	7218	US-08-232-463-14	Sequence 14, Appl
7	55.4	1.1	8438	US-07-945-283-1	Sequence 1, Appl
8	51.8	1.0	4403765	US-09-103-840A-2	Sequence 2, Appl
9	50.4	1.0	1496	US-08-843-993-2	Sequence 2, Appl
10	50.4	1.0	1496	US-09-059-520A-2	Sequence 2, Appl
11	50.4	1.0	1496	US-09-334-275-2	Sequence 2, Appl
12	50.4	1.0	30001	US-08-125-468-1	Sequence 1, Appl
13	50.4	1.0	30001	US-08-474-933-1	Sequence 1, Appl
14	49.4	1.0	2744	US-09-071-101-1	Sequence 1, Appl
15	49.4	1.0	2744	US-09-369-618-1	Sequence 1, Appl
16	49.4	1.0	2744	US-09-369-618-1	Sequence 1, Appl
17	49	1.0	4695	US-08-231-193A-57	Sequence 1, Appl
18	49	1.0	4695	US-08-486-273A-57	Sequence 57, Appl
19	49	1.0	4695	US-08-940-086A-57	Sequence 57, Appl
20	49	1.0	4695	US-08-940-086A-57	Sequence 57, Appl
21	49	1.0	4695	US-08-935-105A-57	Sequence 57, Appl
22	49	1.0	4695	US-08-648-797-57	Sequence 57, Appl
23	48.6	1.0	1147	US-08-761-277A-44	Sequence 44, Appl
24	47.6	0.9	1865	US-08-083-948-7	Sequence 7, Appl
25	47.6	0.9	1865	US-08-393-785-7	Sequence 7, Appl
26	47.6	0.9	1865	US-08-475-694-7	Sequence 7, Appl
27	47.6	0.9	1865	US-08-712-057-7	Sequence 7, Appl

c 28	47.6	0.9	2374	4	US-09-347-801-3	Sequence 3, Appl
c 29	47.6	0.9	2721	6	US-08-775-009-36	Sequence 36, Appl
c 30	47.6	0.9	3507	2	US-08-998-416-861	Sequence 861, Appl
c 31	46.8	0.9	614	4	US-09-178-109-3	Sequence 3, Appl
c 32	46.8	0.9	2064	4	US-09-142-791A-3	Sequence 3, Appl
c 33	46.8	0.9	2072	4	US-09-142-791A-3	Sequence 3, Appl
c 34	46.8	0.9	2104	4	US-09-142-791A-3	Sequence 3, Appl
c 35	46.8	0.9	2104	4	US-09-142-791A-3	Sequence 3, Appl
c 36	46.8	0.9	2121	4	US-09-178-109-1	Sequence 1, Appl
c 37	46.6	0.9	2353	2	US-08-922-068A-01	Sequence 1, Appl
c 38	46.4	0.9	43280	2	US-08-804-227C-1	Sequence 1, Appl
c 39	46.2	0.9	1761	4	US-09-504-358-19	Sequence 19, Appl
c 40	46.2	0.9	1761	4	US-09-504-358-19	Sequence 19, Appl
c 41	46.2	0.9	2690	4	US-09-029-755C-1	Sequence 1, Appl
c 42	46.2	0.9	5117	3	US-08-854-585-1	Sequence 1, Appl
c 43	46.2	0.9	5117	3	US-08-854-585-1	Sequence 1, Appl
c 44	46.2	0.9	11471	5	US-09-504-358-16	Sequence 16, Appl
c 45	46.2	0.9	11471	4	US-09-954-314-16	Sequence 16, Appl

#### ALIGNMENTS

RESULT 1	US-09-907-843-3	Sequence 3, Application US/09907843
Patent No. 6440739		
GENERAL INFORMATION:		
APPLICANT: C. Frank Bennett		
APPLICANT: Susan M. Freiler		
FILE OF INVENTION: ANTISENSE MODULATION OF GLIOMA-ASSOCIATED ONCOGENE-2 EXPRESSION		
CURRENT APPLICATION NUMBER: US/09/907.843		
CURRENT FILING DATE: 2001-07-17		
NUMBER OF SEQ ID NOS: 87		
SEQ ID NO 3		
LENGTH: 4960		
TYPE: DNA		
ORGANISM: Homo sapiens		
FEATURE:		
NAME/KEY: CDS		
LOCATION: (572)...(4348)		
US-09-907-843-3		
Query Match	14.0%;	Score 707; DB 4; Length 4960;
Best Local Similarity	61.4%;	Pred. 1.9e-156;
Matches 1441; Conservative	0;	Mismatches 780; Indels 126; Gaps 14;
QY 1104	CGTTCGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	1163
DB 445	CTTCCCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	504
QY 1164	AGGTTCCGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	1223
DB 505	GGGTTCCGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	564
QY 1224	GAGGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	1278
DB 565	GGGCGGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	624
QY 1279	CGTTCGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	1334
DB 625	CGTTCGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	684
QY 1335	TGACCGGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	1391
DB 685	CAACCGGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	744
QY 1392	AGGCGGCGGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	1435
DB 745	GGGCGGCGGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	804
QY 1436	TCAGGCGGCGGCGGCGGCGGCGTCTCTCCACATGATGATGAGATCCTAAGCGAGCAAGAGCTT	1463



; CURRENT FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 88  
 ; SEQ ID NO 3  
 ; LENGTH: 3600  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (79)....(3399)  
 ; US-09-657-042A-3

Query Match 6.5%; Score 329.4; DB 4; Length 3600;  
 Best Local Similarity 70.6%; Pred. No. 7.9e-68;  
 Matches 454; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

QY 1372 GATGAAGACCTCCCGCCAGGGGCTCGGGGCGACACAGAAAGCCCGAGAGAACACC 1431  
 DB 664 GAAGGTGATATGTCTCCCACTCCACGATACAGGATCCCTGTTGGGGATCTG 723  
 QY 1432 CTGTCAAGGAGAGAGGACAAAGATGAAGCAAGAGGCTGAAGTCAATCATATGAG 1491  
 DB 724 GATGGGGGAGAGACCTCGAGAGAGAGA---GAAGCTGAGCTGATCTGTATGAA 780  
 QY 1492 ACAAACTGCCACTGGGAAGGCTGCGGAGAGTTCGACACCCAGAGAGAGCTTGTGCAC 1551  
 DB 781 ACTGACTGCGCTGGGTGAGTGTCTGACGACAGAAATTTGACTCCCAAGAGAGCTGTGCAC 840  
 QY 1552 CATATTAATACGACCATATTCATGAGAGAGAGAGAGAGTGTCTGTGCAAGTGTGCAC 1611  
 DB 841 CACATCAACAGCGAGACATCCACGGGAGCGGAGAGAGTTCGTGTGCACATGGGGGCG 900  
 QY 1612 TGCTCAAGAGAGAGAAACCTTCACAAAGCCAGTATATGTGTGTATGATATGAGAGA 1671  
 DB 901 TGCTCAAGAGAGAGTGAAGGCTTCACAAAGCCAGTATATGTGTGTATGATATGAGAGA 960  
 QY 1672 CACAGGGGAGAGAGGCTTCACAAAGTTCATTTGAAGGTTGCAAAAGGCTTACTGAGA 1731  
 DB 961 CACACTGGGAGAGAGGCTTCACAAAGTTCATTTGAAGGTTGCAAAAGGCTTACTGAGA 1020  
 QY 1732 CTGAAAACCTTGAAGACACTTGCATCTGACACTGAGAGAGAAACCATATGCTGTGAG 1791  
 DB 1021 CTGAAAACCTTGAAGACACTTGCATCTGACACTGAGAGAGAAACCATATGCTGTGAG 1080  
 QY 1792 CAGGAAGTGTGCAACAAGGCTTCTCAAAATGCTGTATGATGCGGCCCAACCAACAG 1851  
 DB 1081 CAGAGAGGCTGCAATGAGGCTTCTCAAAATGCTGTATGATGCGGCCCAACCAACAG 1140  
 QY 1852 AGGCAATTCATGAGAAACCATATGCTGTGCAAAATGCTGTATGATGCGGCCCAACCAACAG 1911  
 DB 1141 ACCCAATTCATGAGAAACCATATGCTGTGCAAAATGCTGTATGATGCGGCCCAACCAACAG 1200  
 QY 1912 GACCAAGGCTTCTCCGGAAGACATGAGAGACATGAGAGAGGCTCAATGCTATGTCAC 1971  
 DB 1201 GATCTTGAAGGCTTCTCCGGAAGACATGAGAGACATGAGAGAGGCTCAATGCTATGTCAC 1260  
 QY 1972 AAGAGAGGAGAGAGACATGCTCTCGGCGCCACCCCGCA 2014  
 DB 1261 AAGAGAGGAGAGAGACATGCTCTCGGCGCCACCCCGCA 1303

RESULT 3  
 ; US-09-234-332-5  
 ; Sequence 5, Application US/09234332A  
 ; Patent No. 6087168  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cedars-Sinai Medical Center  
 ; APPLICANT: Michael F. Levesque, M.D.  
 ; APPLICANT: Thomas Neuman, Ph.D.  
 ; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO  
 ; TITLE OF INVENTION: NEURONS; TRANSDIFFERENTIATION OF EPIDERMAL CELLS  
 ; FILE REFERENCE: P07 41494  
 ; CURRENT APPLICATION NUMBER: US/09/234,332A  
 ; CURRENT FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 5  
 ; LENGTH: 3138  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: gene  
 ; LOCATION: (0)....(0)  
 ; OTHER INFORMATION: Zic 1 Protein gene; Genbank Accession D76435  
 ; US-09-234-332-5

Query Match 1.8%; Score 90.8; DB 3; Length 3138;  
 Best Local Similarity 54.9%; Pred. No. 8.1e-12;  
 Matches 226; Conservative 0; Mismatches 177; Indels 9; Gaps 2;

QY 1530 CACCCAGAGAGCTTGTGACACCATATTAATTAACGACATATTCATGAGAGAGAGA 1589  
 DB 1524 CATGACAGAGTATGTTACGACGACGTCACCGTGGAGCAGTATGAGTGGCCGAGAGAGTAA 1583  
 QY 1590 GTTCGTGTGAGTGTGAGTGTGACGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGATAT 1649  
 DB 1584 TCACATCTGCTTGTGAGAGAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1643  
 QY 1650 GTTGTGATGATGATGAG 1709  
 DB 1644 ACTGTTAAGCAGATCCGCTGTCACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1703  
 QY 1710 TTGCACAAAGGCTTCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1769  
 DB 1704 CTGTGGCAAGGCTTCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1763  
 QY 1770 AGAGAAACCATATGCTGTGAG 1829  
 DB 1764 GGAGAGAGGCTTCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1823  
 QY 1830 TCGCGCAAG 1889  
 DB 1824 CCGCAAG 1877  
 QY 1890 AGGCTGACATGAGGCTTACACAG 1941  
 DB 1878 ---GTGCGAG 1926

RESULT 4  
 ; US-09-172-045-1  
 ; Sequence 1, Application US/09172045  
 ; Patent No. 6277594  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mikoshiba, Katsuniko  
 ; APPLICANT: Aruga, Jun  
 ; APPLICANT: Nagata, Takeharu  
 ; APPLICANT: Nakata, Katsunori  
 ; TITLE OF INVENTION: Neurogenesis Inducing Gene  
 ; FILE REFERENCE: Htraki-03497  
 ; CURRENT APPLICATION NUMBER: US/09/172,045  
 ; EARLIER FILING DATE: 1998-10-08  
 ; EARLIER APPLICATION NUMBER: JP98/86979  
 ; EARLIER FILING DATE: 1998-03-31  
 ; EARLIER APPLICATION NUMBER: JP98/121456  
 ; EARLIER FILING DATE: 1998-04-30  
 ; NUMBER OF SEQ ID NOS: 40  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2364  
 ; TYPE: DNA  
 ; ORGANISM: Xenopus laevis  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (138)....(1463)  
 ; US-09-172-045-1

Query Match 1.6%; Score 83.2; DB 4; Length 2364;  
 Best Local Similarity 51.5%; Pred. No. 4.4e-10;  
 Matches 271; Conservative 0; Mismatches 243; Indels 12; Gaps 3;

QY 1506 GGAAGGCTGCGGAGGAGTTCGACCCAGAGAGCTTGTGACCATATATAACGA 1565  
 DB 845 GAAACCTGTGACAGACATTAGACATGATGACTGTTACACATATGACAATGCA 904  
 QY 1566 CCATATTCATG--AGAGAAAGAGAGTCTGTGACAGTGGCTGAGACTGCTCAAGGA 1622  
 DB 905 ACATATTTGGGGTCCAGAAACAAATATACATATATGCTACTGAGAGATGTCGAGGG 964  
 QY 1623 GCAGAAACCTTCAAGCCCATATATGTTGGTACTGATATGAAACACACGCGGGA 1682  
 DB 965 AGCTAAATCTTTTAAAGCAATATTAATGATGATGATGATGATGATGATGATGATG 1024  
 QY 1683 GAGGCTTCACAAATGCACTTTTGAAGGTTGACAAAGGCTTACTGAGACTAGAAAATT 1742  
 DB 1025 AAAACCTTTCATGCCCCCTTCCCTGATGTTGGAAATCTTTCGACGCTTCAGAAAATCT 1084  
 QY 1743 GAAACACACTGAGATCTCACACTGAGAGAAACCATGCTGTGAGACAGAGTTG 1802  
 DB 1085 CAGATATCCAAAGAAACATCATAGGTGAGAGGATTCAGTGTGAGTTGAAAGCTG 1144  
 QY 1803 CACAAGGCTTCTCAATGCTCTGATGCGCCCAACACACAAACAGACGATTCGAA 1862  
 DB 1145 CATGAAAGGTTTGCACACAGAGGCA---CAGGAAAAACATATGCAATGTCACACGTC 1201  
 QY 1863 TGAGAAACCATATGTTGCAAAATCCAGAGCTGACATTAAGCTTACAGACGCAAGCTC 1922  
 DB 1202 ACATTAAGCATATCTGCA--AGTGTGATTAATCTCTACACTCACGCCAGCTC 1255  
 QY 1923 CCTCGGAAACATGTAACAGACAGTGCAGAGGCTCATGTCACCAAGCAAGCAGG 1982  
 DB 1256 CCTAAGAAAGCAATGAGGTTCAATGATCAAGAGGCTGATCTTCCCTCGTCCAG 1315  
 QY 1983 AGGGGACATCATCTCGGCGCCGACCCCGAGAGATTCCGCGACG 2028  
 DB 1316 CTCAGGGTACGAATCTCTACCCACAGCAATGTTTCTGTCACAC 1361

RESULT 5  
 US-08-232-463-14  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHEIFLINGER, F.  
 ; APPLICANT: FALKNER, F. G.  
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232.463  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/935.313  
 ; FILING DATE:  
 ; APPLICATION NUMBER: EP 91 114 300.6  
 ; FILING DATE: 26-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7218 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 CLONE: pTZ9pt-F15  
 US-08-232-463-14

Query Match 1.4%; Score 68.6; DB 1; Length 7218;  
 Best Local Similarity 5.0%; Pred. No. 1.8e-06;  
 Matches 20; Conservative 231; Mismatches 150; Indels 0; Gaps 0;

QY 372 GACGCTTTTCCATGAGACCCAGGAATGTTACATGAGCCCACTACCACTCTCTCA 431  
 DB 1052 GAGGAGCTTCCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1111  
 QY 432 TCTTTCCCTCCCTCCATCCCTGATGATGATGATGATGATGATGATGATGATGATG 491  
 DB 1112 YY 1171  
 QY 492 CCATTAAGATCATCTCCGATTCCTCATGATGATGATGATGATGATGATGATGATG 551  
 DB 1172 YY 1231  
 QY 552 TACGATCCGAGCTGCGCTTATAGATCTCCGACACCGGAGCCGCTGCTGTC 611  
 DB 1232 YY 1291  
 QY 612 CGAGTTCCTCCCTCCGATCCCTCATGATGATGATGATGATGATGATGATGATG 671  
 DB 1292 YY 1351  
 QY 672 CTTCACAGCAGCCATCCCTCATGATGATGATGATGATGATGATGATGATGATG 731  
 DB 1352 YY 1411  
 QY 732 TGCGCCCATGACAGAGTACGCCAGCAGATTAATTAATC 772  
 DB 1412 YYYYYYYYYYYYYYYYYYGTACCAATATCTCTATC 1452

RESULT 6  
 US-08-232-463-14/c  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHEIFLINGER, F.  
 ; APPLICANT: FALKNER, F. G.  
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

```

```

Query Match 1.2%; Score 61; DB 1; Length 7218;
Best Local Similarity 2.6%; Pred. No. 0.00011;
Matches 10; Conservative 230; Mismatches 145; Indels 0; Gaps 0;

```

```

QY 1317 TGCAGTACGAGCAGCTGTCAGCCGATGACACAGAGGTCACATCAACCCGATGA 1376
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378
QY 1377 AGACCTCCAGCCAGGCGCTGGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1436
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318
QY 1437 CAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1496
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258
QY 1497 CTGCCACTGGGAGGCTGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1556
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198
QY 1557 AATATACGACCATATTCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1616
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138
QY 1617 AAGGAGCAGAAACCCCTCAAGCCGATATGTTGTTGATGATGATGATGATGATGAT 1676
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1137 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078
QY 1677 GGGGAGAGGCTCACAATGACT 1701
| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1077 RRRRRRRRRRRATCGCAGCTCCT 1053

```

```

RESULT 7
US-07-945-283-1/c
; Sequence 1, Application us/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The Ep0 and LfT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria

```

```

STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P.
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8438 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudorabies virus
FEATURE:
NAME/KEY: CDS
LOCATION: 622..6495
FEATURE:
NAME/KEY: variation
LOCATION: replace(1099, "g")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1267, "t")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1381, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1566, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(7010, "g")
US-07-945-283-1

```

```

Query Match 1.1%; Score 55.4; DB 1; Length 8438;
Best Local Similarity 45.9%; Pred. No. 0.0025;
Matches 262; Conservative 0; Mismatches 306; Indels 3; Gaps 2;

```

```

QY 2604 CGCCAGCAGCATCAGCTCGCTACCTGAGCAGCGCGCTCTCAGGAGATCGCCCTG 2663
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 6155 CCGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6096
QY 2664 CTTCTCAGCGCGCTCCAGGAGGCGTCCAGGAGGCGGCGGCGGCGGCGGCGGCG 2723
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 6095 CTTCCAGCAGCTCCCGCGCGCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCT 6036
QY 2724 C--GTGGCGCATCTCAGACCCCATCTCCAGCAGGCTGCGCGCGCTCCAGCAGG 2781
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 6035 CCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5976
QY 2782 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2841
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 5975 CCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5916
QY 2842 AAGGCCAGTACGAGGCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2901
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 5915 ACCCCAACTCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 5856

```









Db 8514 GACTGCGGTTGAGAGCGCCGACACCGCTTCAACGGGANTTGAAGCGCGGTG 8573  
 QY 3083 GGACAGGCTCCGAGGGGCTGGCCCTGCTC 3112  
 Db 8574 GGAAGCTGGCGCTGCCCAACACGACGCG 8603

## RESULT 14

US-09-071-101-1  
 ; Sequence 1, Application US/09071101  
 ; Patent No. 6013503  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lok, SI  
 ; APPLICANT: Jaspers, Stephen R.  
 ; TITLE OF INVENTION: HUMAN PROHORMONE CONVERTASE 4  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Zymogenetics, Inc.  
 ; STREET: 1201 Eastlake Avenue East  
 ; CITY: Seattle  
 ; STATE: WA  
 ; COUNTRY: USA  
 ; ZIP: 98102  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/071,101  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Parker, Gary E.  
 ; REGISTRATION NUMBER: 31,648  
 ; REFERENCE/DOCKET NUMBER: 97-05  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 206-442-6673  
 ; TELEFAX: 206-442-6678  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2744 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: Coding Sequence  
 ; LOCATION: 61...2325  
 ; OTHER INFORMATION:  
 ; US-09-071-101-1

Query Match 1.0%; Score 49.4; DB 3; Length 2744;  
 Best Local Similarity 44.1%; Pred. No. 0.04;  
 Matches 206; Conservative 0; Mismatches 261; Indels 0; Gaps 0;

QY 2582 TGCTCAACAGAGGACAGAGCGCCGACATCACTGAGGCTTACCTGAGCAGCGCC 2641  
 Db 908 TGACCAAGGGCCGCGGGGCTGGGACGCTTTCATCTGGGCTTGGGCAAGCGGGCC 967  
 QY 2642 GCTCCTCAGGAGTCTGCGCTTCTTCAAGCGCGCTTCAAGCGGCGTCAAGCGCG 2701  
 Db 968 TGCACTACGACACTGCACTGCGAGCGGTACACAAACAGCATTCACAGCTTTCGCTG 1027  
 QY 2702 AGGGCCGCGCGCAGAAAGTGAAGGTGGCCGACCTCTTACGACCCCATCTCCACGAGGCT 2761  
 Db 1028 GCACACACACCCAGCAAGGCGCGGTGCGGTGTAACAGCGAAGCGCTGGCTTCCACCTCA 1087  
 QY 2762 CGCGCCCTCAGGAGGAGCGAGCGAGCGGCGCTGCCAGCTGCTCAGCCTCAAGC 2821

Db 1088 CCACCACTTACAGAGCGCGCTGGCCACCGACCCCAATGCTACCAAGGACCTGCATC 1147  
 QY 2822 CCGCCAGCAGTACCGCTCTCAAGGCCAAGTACCGGCTGCGACAGAGGCGCCGCCGA 2881  
 Db 1148 ACGGTGACACAGCAGCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207  
 QY 2882 CCGCCCTGCGCAACATGAGAGGATGAGCGCTGAGACGCGCTGCGCTGCGGAGATC 2941  
 Db 1208 TCGCCCTTACGCTGAGAGGCCAACCCGTTCTGACTGAGAGACATGACGACCTGTGG 1267  
 QY 2942 CCGTCAAGCCTGAGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3001  
 Db 1268 TCGCGCGCTTCAAGCGCGCGCACCTGCGAGGCGGAGAGCTGAGAGCAACAGCGCGTGG 1327  
 QY 3002 GAGCCACGCTACAGGCGCGCGCCACCTGCGACCGCGACGATGCGCTG 3048  
 Db 1328 GCCAAGTACGATCACTACGATGATGAGGCTGCTGAGCGCGGCGCTG 1374

## RESULT 15

US-09-369-618-1  
 ; Sequence 1, Application US/09369618  
 ; Patent No. 6100041  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lok, SI  
 ; APPLICANT: Jaspers, Stephen R.  
 ; TITLE OF INVENTION: HUMAN PROHORMONE CONVERTASE 4  
 ; FILE REFERENCE: 97-05D2  
 ; CURRENT APPLICATION NUMBER: US/09/369,618  
 ; CURRENT FILING DATE: 1999-08-06  
 ; EARLIER APPLICATION NUMBER: US 09/071,101  
 ; EARLIER FILING DATE: 1998-05-01  
 ; EARLIER APPLICATION NUMBER: US 60/044,015  
 ; EARLIER FILING DATE: 1998-05-06  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO: 1  
 ; LENGTH: 2744  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (61)...(2325)  
 ; US-09-369-618-1

Query Match 1.0%; Score 49.4; DB 3; Length 2744;  
 Best Local Similarity 44.1%; Pred. No. 0.04;  
 Matches 206; Conservative 0; Mismatches 261; Indels 0; Gaps 0;

QY 2582 TGCTCAACAGAGGACAGAGCGCCGACATCACTGAGGCTTACCTGAGCAGCGCC 2641  
 Db 908 TGACCAAGGGCCGCGGGGCTGGGACGCTTTCATCTGGGCTTGGGCAAGCGGGCC 967  
 QY 2642 GCTCCTCAGGAGTCTGCGCTTCTTCAAGCGCGCTTCAAGCGGCGTCAAGCGCG 2701  
 Db 968 TGCACTACGACACTGCACTGCGAGCGGTACACAAACAGCATTCACAGCTTTCGCTG 1027  
 QY 2702 AGGGCCGCGCGCAGAAAGTGAAGGTGGCCGACCTCTTACGACCCCATCTCCACGAGGCT 2761  
 Db 1028 GCACACACACCCAGCAAGGCGCGGTGCGGTGTAACAGCGAAGCGCTGGCTTCCACCTCA 1087  
 QY 2762 CGCGCCCTCAGGAGGAGCGAGCGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 2821  
 Db 1088 CCACCACTTACAGAGCGCGCTGGGCGACCGCCAGATCTGCTACCAAGGACCTGCATC 1147  
 QY 2822 CCGCCAGCAGTACCGCTTCAAGGCCAAGTACGCGGCTGCCACAGAGGCGCGCGCGCA 2881  
 Db 1148 ACGGTGACACAGCAGCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207  
 QY 2882 CGCGCCCTCAGGAGGAGGATGAGCGCTGAGACGCGCTGCGCTGCTGCTGCTGCTGCTG 2941  
 Db 1208 TCGCCCTTACGCTGAGAGGCCAACCCGTTCTGACTGAGAGACATGACGACCTGTGG 1267



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2002, 19:51:21 ; Search time 8366 Seconds  
(without alignments) 17584.829 Million cell updates/sec

Title: US-09-910-185-3

Perfect score: 5055  
Sequence: 1 cgatactacgtggcatttt.....accctcttttaaaaaaa 5055

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.ov:\*  
22: em.pat:\*  
23: em.ph:\*  
24: em.pl:\*  
25: em.ro:\*  
26: em.sts:\*  
27: em.vi:\*  
28: em.yl:\*  
29: em.yl:\*  
30: em.htg.hum:\*  
31: em.htg.in:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.rtd:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

Pred. NO. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5055	100.0	5055	9 HUMGLI3A	M57609 Human DNA-b
2	4732	93.6	4743	10 HSA250408	AF250408 Homo sapi
3	3696.6	73.1	5113	10 HSMGLI3GEN	X85255 M.musculus
4	2758.6	54.6	4962	5 AF231112	AF231112 Coturnix
5	2695.6	53.1	5025	5 AF231111	AF231111 Coturnix
6	2685.8	53.1	4632	5 AF222990	AF222990 Gallus ga
7	2557.6	50.6	18927	5 AC005028	AC005028 Homo sapi
8	2115.8	41.9	6695	2 XLU02461	U42461 Xenopus lae
9	1625.2	32.2	132169	2 AC068502	AC068502 Mus muscu
10	1625.2	32.2	175144	2 AF303886	AF303886 Gallus ga
11	1581.4	31.3	2175	5 AC128489	AC128489 Rattus no
12	1506.8	29.8	186926	2 AC128489	AC128489 Rattus no
13	864.8	17.1	5849	10 MMGLI2	X99104 M.musculus
14	796.2	15.8	4517	5 AF109923	AF109923 Xenopus l
15	789.2	15.6	4641	5 XLU04262	U42462 Xenopus lae
16	770.8	15.2	2525	5 AF022818	AF022818 Gallus ga
17	756	15.0	6070	5 AF085746	AF085746 Danio rer
18	752.2	14.9	3678	5 AB007298	AB007298 Homo sapi
19	752.2	14.9	4909	5 AB007296	AB007296 Homo sapi
20	720.6	14.3	4156	5 AF348156	AF348156 Rattus no
21	707	14.0	3729	5 AB007297	AB007297 Homo sapi
22	707	14.0	4960	5 AB007295	AB007295 Homo sapi
23	699	13.8	1989	5 GGU60763	GGU60763 Rattus no
24	678.6	13.4	16964	2 AC096345	AC096345 Rattus no
25	610.2	12.1	855	5 AF316110	AF316110 Notophtha
26	541.6	10.7	2105	9 HUWTHP1	D14827 Human mRNA
27	496.4	8.5	186826	2 AC128489	AC128489 Human mRNA
28	431.2	7.2	1934	5 GGU60762	GGU60762 Gallus gall
29	363.8	7.1	4081	5 XLU057454	U57454 Xenopus lae
30	360.2	6.6	3686	10 AB025922	AB025922 Mus muscu
31	334.8	6.6	4644	3 BFL252244	AF252244 Branchios
32	333.4	6.6	3648	10 AF026305	AF026305 Mus muscu
33	331.6	6.5	3587	9 HSGLI	BC013000 Homo sapi
34	329.4	6.5	3600	9 HSGLI	X07384 Human mRNA
35	329.4	6.0	155888	9 AC016764	AC016764 Homo sapi
36	325.4	5.1	148658	2 AC095699	AC095699 Rattus no
37	258.6	5.1	329	10 AB073718	AB073718 Rattus no
38	254.4	5.0	4902	3 DMCID	X54360 Drosophila
39	254.4	4.9	266344	9 AC005158	AC005158 Homo sapi
40	246.2	4.8	1937	3 AF091245	AF091245 Junonia c
41	242.6	4.8	1937	3 AC096345	AC096345 Rattus no
42	222.2	4.4	16964	2 AC109271	AC109271 Mus muscu
43	218.6	4.3	198951	2 AC122287	AC122287 Mus muscu
44	218.6	4.3	236215	9 AC005026	AC005026 Homo sapi
45	218	4.3	123526	9 AC005026	AC005026 Homo sapi

## ALIGNMENTS

RESULT 1  
LOCUS HUMGLI3A  
DEFINITION Human DNA-binding protein (GLI3) mRNA, complete cds.  
ACCESSION M57609.1  
VERSION M57609.1 GI:183247  
KEYWORDS SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
GLI3 encodes a 190-kilodalton protein with multiple regions of GLI similarity

JOURNAL Mol. Cell. Biol. 10 (10), 5408-5415 (1990)  
 MEDLINE 90377231  
 PUBMED 2118997  
 FEATURES  
 SOURCE

Location/Qualifiers

1. 5055  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="7p13"  
 /cell\_type="xenograft D245MG"  
 /cell\_line="g110blastoma multiforme"  
 1. 5055  
 /gene="GLI3"  
 <1. 5046  
 /gene="GLI3"  
 /product="DNA-binding protein"  
 /note="G00-119-990"  
 55. 4845  
 /gene="GLI3"  
 /codon\_start=1  
 /product="DNA-binding protein"  
 /protein\_id="AAA52564.1"  
 /db\_xref="GI:183248"  
 /db\_xref="GDB:600-119-990"  
 /translation="MEOSHSSTTEKKKVENISVCKSTRDVESEKAVASSTTSENDE  
 SPQGYHREBRALTMOPONOVGLSKVEEPTSDSRASLKEIKHISLPHVAEPV  
 PYRGTFVADPNRNGMEPHHHPHLPFAHPVPIDRHHGHHVDPSPIDPHMTS  
 ALSSPTYPDLPEFIRISPHRNPAASESPSPIDPNWDRYRSTRHSSPSLMTSA  
 TRGLSPDAPAGVAPAYYHOMALNGORSPVDITPSAATAGCAIHEVYLAHSA  
 TRFSPRLSARSPRRKRLISPLSDHSPDLOMTIRTPNSVLITLNNRSSASASQY  
 GHLSAISPALSFTSSAPVSLMHOOILSRQSLASAGHSPILNHRSPSSASQY  
 IGLPVLNAPVOSGSPESQNPPTSESAVSSDTGMHNRKIRKIPDEDLPPGAG  
 QOODECTIVIKKEGDDKDESKOPEVETYNHMGARCFEDYEOELVHINDHIG  
 EKEFVCKMIDCSREOKPFAQOYMLVVMRRHNGEKPHKCTFEQCTKAVSRLENLKT  
 LRSTGEKPVYCEHGEONKFAQNSADRAKHONRTHSEKRYVCKIKGCTKRYTDP  
 RKHKVYHGEPAHVKRQKRDHPRPPSPDSSGSHOSRSGRSPGAGALGEBODLSNT  
 TSRECELOVYTAKEKPMTSOPSPGSSQSOOPISTSVSGLELPLDGGSLG  
 LSAIDETPIMDSITATLALOAARNGCTMMHVKLERIKOVNGFPLRNPILP  
 PKAVAPPLINGTOSNNICSLGPMILPGRSDLGVDVMTMLNLRDSSASTIS  
 LPLSLSLTAPAOYRIKRYKAATGEGPEPPLPMEMMSLKTALALGDLLEFVALP  
 AYLSSRSISGSPSSRSSEASOAEGRQNVSDVDTSTDAERSSPAOSQD  
 VHPARRCSGAGHGTGRRLQPHDALGHRASDVRVSGEGLALPVRFRSSALSP  
 NPPMATSAKRSILVONATRRPEGOSRNFSPSPSTENVTLTSLMDADNL  
 EDPLEDDVOYLNSONAGYEQHPSPALDDSKVPHPGDPDPAPIGIDSHAGNDR  
 EQPCPESSKTDLPIONNEVSSGADLSKLGKGPAPVORTARFGCMGVYHONP  
 LRSGPAGYOTLIGENSNPYGGPEHMLHNSPGSGTGNAPORCKAPQYCKLNRBP  
 VAGALDGAAGAIQASKLSTPMOSGGOLNGUPVAPNESAGVMGMOMDVGQ  
 GYLHQLDGSOMHPGAGRPGOMLQOITATSHINITYCPESCLPGAGMGSPSSLA  
 VVRGXPQASRFGSRQAMPRLSLALQSLDSTQCRVNIKMEMGQHPILCSNL  
 ONYSGOPYOYVGFESQDTRKASFTISDMSCLQSTAKNSLISPGANOVYSTVDSL  
 DSHDLGVOIDEDPATIDGDHSSLSMGAISP5IIQNLSSSSRLTTPPASLFPVAVH  
 EHHQHEIRGHEFFADLPISGRKOIPCSYALGFRKKRLQETLEINRS"

BASE COUNT 1243 a 1591 c 1290 g 931 t  
 ORIGIN

Query Match 100.0%; Score 5055; DB 9; Length 5055;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGATCTAGCTGGGCAATTTTGGTCGAAGAGAGCTGAATATGAGAGACATCATGAG 60  
 Db 1 CGATCTAGCTGGGCAATTTTGGTCGAAGAGAGCTGAATATGAGAGACATCATGAG 60  
 QY 1 CGATCTAGCTGGGCAATTTTGGTCGAAGAGAGCTGAATATGAGAGACATCATGAG 60  
 61 GCCCAGTCCACAGCTCCACGACACACTGAAAAAAGTTGAGAAATTCATAGTGAAG 120  
 Db 61 GCCCAGTCCACAGCTCCACGACACACTGAAAAAAGTTGAGAAATTCATAGTGAAG 120  
 QY 61 GCCCAGTCCACAGCTCCACGACACACTGAAAAAAGTTGAGAAATTCATAGTGAAG 120  
 121 TGTCTCACTGCAAGATGTGAGCAGAAAGCGTTGCTCCACAGACACACTTAAATGAG 180  
 Db 121 TGTCTCACTGCAAGATGTGAGCAGAAAGCGTTGCTCCACAGACACACTTAAATGAG 180  
 QY 121 TGTCTCACTGCAAGATGTGAGCAGAAAGCGTTGCTCCACAGACACACTTAAATGAG 180  
 181 GATGAAAGTCCGAGACACACTTATCAGAGAGAGAAAGCAATCAGATGACACCA 240  
 Db 181 GATGAAAGTCCGAGACACACTTATCAGAGAGAGAAAGCAATCAGATGACACCA 240  
 QY 181 GATGAAAGTCCGAGACACACTTATCAGAGAGAGAAAGCAATCAGATGACACCA 240

QY 241 CAGAAATGTCAGAGGAGCTCAGCAAGTCAAGTGAAGGACCTTCAACATGAGTACGAGAG 300  
 Db 241 CAGAAATGTCAGAGGAGCTCAGCAAGTCAAGTGAAGGACCTTCAACATGAGTACGAGAG 300  
 QY 301 GCCCATATGATCAAGAAAAAGATCCATGGTCCCTGACACAGCTGGCGAGCCCTGTG 360  
 Db 301 GCCCATATGATCAAGAAAAAGATCCATGGTCCCTGACACAGCTGGCGAGCCCTGTG 360  
 QY 361 CCGTACCGCGGAGGAGTGTGGCATGGAGCCGAGAAATGTTACTATGAGCCCATAC 420  
 Db 361 CCGTACCGCGGAGGAGTGTGGCATGGAGCCGAGAAATGTTACTATGAGCCCATAC 420  
 QY 421 CACCTCCATCTCTTCCCTGCTCCATCCATCCCTGATGATGATGATGATGATGATGAT 480  
 Db 421 CACCTCCATCTCTTCCCTGCTCCATCCCTGATGATGATGATGATGATGATGATGAT 480  
 QY 481 GAGGCGCTTACCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 Db 481 GAGGCGCTTACCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 QY 541 TCTAGTACGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 Db 541 TCTAGTACGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 QY 601 GCTGCTGCTTCCGAGTCTCCCTTCCGAGCTGCTCCATGATGATGATGATGATGATGAT 660  
 Db 601 GCTGCTGCTTCCGAGTCTCCCTTCCGAGCTGCTCCATGATGATGATGATGATGATGAT 660  
 QY 661 TATATCGCTCTTCCGAGTCTCCCTTCCGAGCTGCTCCATGATGATGATGATGATGAT 720  
 Db 661 TATATCGCTCTTCCGAGTCTCCCTTCCGAGCTGCTCCATGATGATGATGATGATGAT 720  
 QY 721 ACCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 Db 721 ACCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 QY 781 CTGCTACTGCGCAGCGAGCGCCCTATGACAGATATATTCCTCAGCTGCGACCGCG 840  
 Db 781 CTGCTACTGCGCAGCGAGCGCCCTATGACAGATATATTCCTCAGCTGCGACCGCG 840  
 QY 841 ACGGGGCGCATCCACATGAGAAATATCTCATGATGATGATGATGATGATGATGATGAT 900  
 Db 841 ACGGGGCGCATCCACATGAGAAATATCTCATGATGATGATGATGATGATGATGATGAT 900  
 QY 901 AGGCTGTCAGCGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 Db 901 AGGCTGTCAGCGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 QY 961 AGCTTTACCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 Db 961 AGCTTTACCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 QY 1021 AATTCCTGTCAGCGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGAT 1080  
 Db 1021 AATTCCTGTCAGCGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGAT 1080  
 QY 1081 AGCCTGCTTGGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGATGATGAT 1140  
 Db 1081 AGCCTGCTTGGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGATGATGAT 1140  
 QY 1141 ATCTTACGCGACACAGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGAT 1200  
 Db 1141 ATCTTACGCGACACAGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGAT 1200  
 QY 1201 CCGGCGCGGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGATGATGAT 1260  
 Db 1201 CCGGCGCGGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGATGATGAT 1260  
 QY 1261 GTCAGAGTCCGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGATGATGAT 1320  
 Db 1261 GTCAGAGTCCGAGCTTTCAGCAAGTGGCTCTTATGATGATGATGATGATGATGATGATGAT 1320

QY 1321 GTGAGCAGCTGCTGAGACCCGATGACACAGAGAGTGCAGATCAAAACCCGATGAGAC 1380  
 |||||||  
 Db 1321 GTGAGCAGCTGCTGAGACCCGATGACACAGAGAGTGCAGATCAAAACCCGATGAGAC 1380  
 |||||||  
 QY 1381 CTCCCGAGCCGAGGAGGCTCGGGGGCAGCAGAGAGCCCAAGAGAACCTTTGTCAG 1440  
 |||||||  
 Db 1381 CTCCCGAGCCGAGGAGGCTCGGGGGCAGCAGAGAGCCCAAGAGAACCTTTGTCAG 1440  
 |||||||  
 QY 1441 GAGAGAGGGGCAAGATGAGAGCAACAGAGAGCTGATCATATGAGAGAACCTGC 1500  
 |||||||  
 Db 1441 GAGAGAGGGGCAAGATGAGAGCAACAGAGAGCTGATCATATGAGAGAACCTGC 1500  
 |||||||  
 QY 1501 CACTGGAGAGGCTGCGGAGGAGTTGACACCCCAAGAGAGCTTGTGCACTATATAAT 1560  
 |||||||  
 Db 1501 CACTGGAGAGGCTGCGGAGGAGTTGACACCCCAAGAGAGCTTGTGCACTATATAAT 1560  
 |||||||  
 QY 1561 AAGGACCATATTTATGAGAGAGAGAGAGAGTGTGTGCAAGGTGCTGCACTCTCAGA 1620  
 |||||||  
 Db 1561 AAGGACCATATTTATGAGAGAGAGAGAGAGTGTGTGCAAGGTGCTGCACTCTCAGA 1620  
 |||||||  
 QY 1621 GAGAGAGAACCTTCAAGGCCAGTATATGTTGATGATGAGAGAACACAGAGGCG 1680  
 |||||||  
 Db 1621 GAGAGAGAACCTTCAAGGCCAGTATATGTTGATGATGAGAGAACACAGAGGCG 1680  
 |||||||  
 QY 1681 GAGAGAGCTCACAATGACATTTTGAAGTTGACAAAGGCTACTGAGACTAGAGAAAC 1740  
 |||||||  
 Db 1681 GAGAGAGCTCACAATGACATTTTGAAGTTGACAAAGGCTACTGAGACTAGAGAAAC 1740  
 |||||||  
 QY 1741 TTGAGAAACACATTTGAGATCTCAGCTGAGAGAGAAACCATAGCTCTGTGACAGAGGT 1800  
 |||||||  
 Db 1741 TTGAGAAACACATTTGAGATCTCAGCTGAGAGAGAAACCATAGCTCTGTGACAGAGGT 1800  
 |||||||  
 QY 1801 TGCACAAAGGCTTTCACAAATGCTGATGAGGCGCAACACCAAAAGAGAGCAATTC 1860  
 |||||||  
 Db 1801 TGCACAAAGGCTTTCACAAATGCTGATGAGGCGCAACACCAAAAGAGAGCAATTC 1860  
 |||||||  
 QY 1861 AATGAGAAACCATATGTTGTCACAAATCCAGGCTGACACTAGGCTTACACAGAGCCAGC 1920  
 |||||||  
 Db 1861 AATGAGAAACCATATGTTGTCACAAATCCAGGCTGACACTAGGCTTACACAGAGCCAGC 1920  
 |||||||  
 QY 1921 TCCCTCCGGAACATGTTGAAGAGAGTGCATGAGCCGACAGAGCTCATGTACCAAGAGCAG 1980  
 |||||||  
 Db 1921 TCCCTCCGGAACATGTTGAAGAGAGTGCATGAGCCGACAGAGCTCATGTACCAAGAGCAG 1980  
 |||||||  
 QY 1981 CGAGGGAGCATCCATCCTCGGCGCCGACCCCGAGAGATTCGGGAGCCATTACAGTCC 2040  
 |||||||  
 Db 1981 CGAGGGAGCATCCATCCTCGGCGCCGACCCCGAGAGATTCGGGAGCCATTACAGTCC 2040  
 |||||||  
 QY 2041 AGGTGCGCTGGCGGACCGACTAGGAGGAGCCCTTGTGAGCAGAGAGACCTCAGCAACACT 2100  
 |||||||  
 Db 2041 AGGTGCGCTGGCGGACCGACTAGGAGGAGCCCTTGTGAGCAGAGAGACCTCAGCAACACT 2100  
 |||||||  
 QY 2101 ACCTCAAAAGGGGAGAGATTCCTCCAGGTGAAAAACCGTCAAGGCAAGAGCCAAATGACA 2160  
 |||||||  
 Db 2101 ACCTCAAAAGGGGAGAGATTCCTCCAGGTGAAAAACCGTCAAGGCAAGAGCCAAATGACA 2160  
 |||||||  
 QY 2161 TCTCAGCAGAGCCCTGGTGTGCTCATGTGAGCAGAGCAACAGTCCCATCAGCAAC 2220  
 |||||||  
 Db 2161 TCTCAGCAGAGCCCTGGTGTGCTCATGTGAGCAGAGCAACAGTCCCATCAGCAAC 2220  
 |||||||  
 QY 2221 TATTTCAACAGTGGGCTCGAGCTTCTCTGACCGATGAGAGTATAGAGAGCTCAGT 2280  
 |||||||  
 Db 2221 TATTTCAACAGTGGGCTCGAGCTTCTCTGACCGATGAGAGTATAGAGAGCTCAGT 2280  
 |||||||  
 QY 2281 GGCATGATGAGAAACCAATGAGTCAACATTTTCCACTGCAACACAGGCTTGTCT 2340  
 |||||||  
 Db 2281 GGCATGATGAGAAACCAATGAGTCAACATTTTCCACTGCAACACAGGCTTGTCT 2340  
 |||||||  
 QY 2341 TTGCAAGCCAGAGAAACCCGCGAGAGCAAAATGATGAGAGCAGCTTAAATAGAGAG 2400  
 |||||||  
 Db 2341 TTGCAAGCCAGAGAAACCCGCGAGAGCAAAATGATGAGAGCAGCTTAAATAGAGAG 2400  
 |||||||  
 QY 2401 CTAAACAAAGTGAATGATGTTTCCGCGACTGAGACCCCATTTCTACCCCTTAAAGCCCT 2460  
 |||||||

Db 2401 CTAAACAAAGTGAATGATGTTTCCGCGACTGAGAACCCCATTTCTACCCCTTAAAGCCCT 2460  
 |||||||  
 QY 2461 GCGGTCTCTCTCTATAGAGAAATGGACACAGTCCAAACACACCTGCAGCTTGGGTGG 2520  
 |||||||  
 Db 2461 GCGGTCTCTCTCTATAGAGAAATGGACACAGTCCAAACACACCTGCAGCTTGGGTGG 2520  
 |||||||  
 QY 2521 CCCATGAGCTTCTCCCGGAGAGGCGCTCTGCGGTGAGAGTACATATGCTGAC 2580  
 |||||||  
 Db 2521 CCCATGAGCTTCTCCCGGAGAGGCGCTCTGCGGTGAGAGTACATATGCTGAC 2580  
 |||||||  
 QY 2581 ATGCTCAACAGAGGAGAGCAGAGCCGACACATCAGTCCGCTTACTTGAAGAGCCG 2640  
 |||||||  
 Db 2581 ATGCTCAACAGAGGAGAGCAGAGCCGACACATCAGTCCGCTTACTTGAAGAGCCG 2640  
 |||||||  
 QY 2641 CGTCTCTCAGGAGTCTCGCCCTGCTTCTCAGAGCCGCTCTCAGAGAGGCTCAGAGCC 2700  
 |||||||  
 Db 2641 CGTCTCTCAGGAGTCTCGCCCTGCTTCTCAGAGCCGCTCTCAGAGAGGCTCAGAGCC 2700  
 |||||||  
 QY 2701 GAGGGCGGCGGAGAGAGTGAAGGCGTGGCGACTCTACAGACCCCATCTCCACAGAGCC 2760  
 |||||||  
 Db 2701 GAGGGCGGCGGAGAGAGTGAAGGCGTGGCGACTCTACAGAGCCCATCTCCACAGAGCC 2760  
 |||||||  
 QY 2761 TCGCGCGCTCCAGAGGAGCCAGAGAGCAGGAGCTGAGCCGCTGACAGCTCAGC 2820  
 |||||||  
 Db 2761 TCGCGCGCTCCAGAGGAGCCAGAGAGCAGGAGCTGAGCCGCTGACAGCTCAGC 2820  
 |||||||  
 QY 2821 CCCGCCAGAGTACCGGCTCAAGGCGCAAGTACGCGGCTGCCACAGAGAGGCGCGCG 2880  
 |||||||  
 Db 2821 CCCGCCAGAGTACCGGCTCAAGGCGCAAGTACGCGGCTGCCACAGAGAGGCGCGCG 2880  
 |||||||  
 QY 2881 AGCCCTGCGCCACATGAGAGAGTGAAGGCTGAGAGAGCCGCTGAGGCTCTGCGGGAT 2940  
 |||||||  
 Db 2881 AGCCCTGCGCCACATGAGAGAGTGAAGGCTGAGAGAGCCGCTGAGGCTCTGCGGGAT 2940  
 |||||||  
 QY 2941 GCCCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 3000  
 |||||||  
 Db 2941 GCCCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 3000  
 |||||||  
 QY 3001 GAGGCCAGAGGCTACGAGGCGGCGCCACCTGACGCGCAGATGCGGTGAGCGCGCG 3060  
 |||||||  
 Db 3001 GAGGCCAGAGGCTACGAGGCGGCGCCACCTGACGCGCAGATGCGGTGAGCGCGCG 3060  
 |||||||  
 QY 3061 AGGAGGGCCAGAGGAGCCGCTGAGAGGCTGAGAGGCGCTGAGGCTGAGGCTGAGGCTGAGG 3120  
 |||||||  
 Db 3061 AGGAGGGCCAGAGGAGCCGCTGAGAGGCTGAGAGGCGCTGAGGCTGAGGCTGAGGCTGAGG 3120  
 |||||||  
 QY 3121 CGCTTACAGAGCTGAGAGCTGAGAGCCCGCGGAGATGAGGCTGAGGCTGAGGCTGAGG 3180  
 |||||||  
 Db 3121 CGCTTACAGAGCTGAGAGCTGAGAGCCCGCGGAGATGAGGCTGAGGCTGAGGCTGAGG 3180  
 |||||||  
 QY 3181 AGTCTGCTGCTTCAAGATTTACAGCGGCGCGGAGAGGCGGCTGAGGCTGAGGCTGAGG 3240  
 |||||||  
 Db 3181 AGTCTGCTGCTTCAAGATTTACAGCGGCGCGGAGAGGCGGCTGAGGCTGAGGCTGAGG 3240  
 |||||||  
 QY 3241 TCCCTGCTGCTTCCAGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 3300  
 |||||||  
 Db 3241 TCCCTGCTGCTTCCAGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 3300  
 |||||||  
 QY 3301 GATGCCAACCTGAGAGATGAGATTTCTGCGGAGAGAGCTGAGATTTAAATTC 3360  
 |||||||  
 Db 3301 GATGCCAACCTGAGAGATGAGATTTCTGCGGAGAGAGCTGAGATTTAAATTC 3360  
 |||||||  
 QY 3361 CAGAACCAAGAGGCTAGAGAGAGCTTCCAGAGGCTTCCGAGAGAGAGAGAGAG 3420  
 |||||||  
 Db 3361 CAGAACCAAGAGGCTAGAGAGAGCTTCCAGAGGCTTCCGAGAGAGAGAGAGAGAG 3420  
 |||||||  
 QY 3421 CCCCAAGGGCGGCTGAGCTTGAAGCGCGCGGCTGAGAGAGAGAGAGAGAGAGAG 3480  
 |||||||  
 Db 3421 CCCCAAGGGCGGCTGAGCTTGAAGCGCGCGGCTGAGAGAGAGAGAGAGAGAGAGAG 3480  
 |||||||  
 QY 3481 TTTCATGCTTGAAG 3540  
 |||||||

Db 3481 TTCCATGCCCTTCGAGAGCCCTGCCCCGAGGAGCAAAACCGACCTCCCATTCAGTGG 3540  
 QY 3541 AACGAAATGAGTCCGGGAAAGGCGGAGCCCTGCTCTCCCAAGCTCAAGTGGGGCGGAG 3600  
 Db 3541 AACGAAATGAGTCCGGGAAAGGCGGAGCCCTGCTCTCCCAAGCTCAAGTGGGGCGGAG 3600  
 QY 3601 CCCGCTGTCGCGAGACTGCGCCCTTTGGTTCCTGCAACGGGATGGTCCACCCGAG 3660  
 Db 3601 CCCGCTGTCGCGAGACTGCGCCCTTTGGTTCCTGCAACGGGATGGTCCACCCGAG 3660  
 QY 3661 AACCCCTTGAGAGGCGGCGCTCTGCGGGGCTATCAGACCCCTCGGGGAGAACAGCAACCC 3720  
 Db 3661 AACCCCTTGAGAGGCGGCGCTCTGCGGGGCTATCAGACCCCTCGGGGAGAACAGCAACCC 3720  
 QY 3721 TACGGGCGCCAGAGACTGATGCTCCACACAGCCCGGAGAGTGGACCGCTGGAAC 3780  
 Db 3721 TACGGGCGCCAGAGACTGATGCTCCACACAGCCCGGAGAGTGGACCGCTGGAAC 3780  
 QY 3781 GCCCTTCATGAACAGCCCTGTAAAGCCCGCAGTATGGAACTGTCTCAACAGGAGCCA 3840  
 Db 3781 GCCCTTCATGAACAGCCCTGTAAAGCCCGCAGTATGGAACTGTCTCAACAGGAGCCA 3840  
 QY 3841 GTGGCCCTGCTGCTACCTGACGCTGCTGTGTGTCGCGGATCAAGCCTCAAGCTGAAG 3900  
 Db 3841 GTGGCCCTGCTGCTACCTGACGCTGCTGTGTGTCGCGGATCAAGCCTCAAGCTGAAG 3900  
 QY 3901 AACACCCCATGCAAGGAGCGGGGCGCAGTGAATTTGCGCTCGCGGTACGCGCAAT 3960  
 Db 3901 AACACCCCATGCAAGGAGCGGGGCGCAGTGAATTTGCGCTCGCGGTACGCGCAAT 3960  
 QY 3961 GAGTCAGCTGGAGAGTGTGATGAGCATGAGAACACAGAGCCAGTGGGAGAGGGGTAC 4020  
 Db 3961 GAGTCAGCTGGAGAGTGTGATGAGCATGAGAACACAGAGCCAGTGGGAGAGGGGTAC 4020  
 QY 4021 CTGGCTCAGCAGCTCTCTGCGGACAGCATGACAGCAGCCGGGGGAGCGCCCGCTCAG 4080  
 Db 4021 CTGGCTCAGCAGCTCTCTGCGGACAGCATGACAGCAGCCGGGGGAGCGCCCGCTCAG 4080  
 QY 4081 CAGATGCTTGGGAGATTTAGTGTCTACATCAATCAATATACCAAGGCGCAGAGAGC 4140  
 Db 4081 CAGATGCTTGGGAGATTTAGTGTCTACATCAATCAATATACCAAGGCGCAGAGAGC 4140  
 QY 4141 TGCCTGCGCAGGGGCTTCACGCGCATGGGACAGCCGCTCAAGCTTGGCAGTTGTCAAGGCG 4200  
 Db 4141 TGCCTGCGCAGGGGCTTCACGCGCATGGGACAGCCGCTCAAGCTTGGCAGTTGTCAAGGCG 4200  
 QY 4201 TACCAAGCATTTGCTGCTTTGGGGGAGCAGAGCCGACAGGCTATGCCAGGAGCAGCCTT 4260  
 Db 4201 TACCAAGCATTTGCTGCTTTGGGGGAGCAGAGCCGACAGGCTATGCCAGGAGCAGCCTT 4260  
 QY 4261 GCTGCTCAGTCAAGAGCATGCTCAGTACAGCAAGTCAAGACCTGAGAGGTGATGATCAAG 4320  
 Db 4261 GCTGCTCAGTCAAGAGCATGCTCAGTACAGCAAGTCAAGACCTGAGAGGTGATGATCAAG 4320  
 QY 4321 ATGGAATGAAAGGAGGAGCCCATCCGCTGCTCTCAATCTCAGAAATTAATCTGCTGAG 4380  
 Db 4321 ATGGAATGAAAGGAGGAGCCCATCCGCTGCTCTCAATCTCAGAAATTAATCTGCTGAG 4380  
 QY 4381 TTCTATGACCAAAACCGTGGGCTTCACTGACGAAAGACAGCAAAAGCTGTTCAATCTCTAT 4440  
 Db 4381 TTCTATGACCAAAACCGTGGGCTTCACTGACGAAAGACAGCAAAAGCTGTTCAATCTCTAT 4440  
 QY 4441 TCAGAGCCAGTGGCTGTGTACAGGAGGAGCCGCAAAACTGAGTACTTTCCCA 4500  
 Db 4441 TCAGAGCCAGTGGCTGTGTACAGGAGGAGCCGCAAAACTGAGTACTTTCCCA 4500  
 QY 4501 GGTGCTAATCAGGTGACAGCAGTGTGACAGCTTCAGAGCCATGACCTGGAAAGGAGTA 4560  
 Db 4501 GGTGCTAATCAGGTGACAGCAGTGTGACAGCTTCAGAGCCATGACCTGGAAAGGAGTA 4560  
 QY 4561 CAGATTGACTTGCATCATATAGAGATGGGAGACCTCCAGCCTGATGTGCGGGGCGC 4620  
 Db 4561 CAGATTGACTTGCATCATATAGAGATGGGAGACCTCCAGCCTGATGTGCGGGGCGC 4620

QY 4621 CTGAGCCCAAGTATCATTCAGAACTTTCCCATAGTCTCCCGCTCACAGCCCTCG 4680  
 Db 4621 CTGAGCCCAAGTATCATTCAGAACTTTCCCATAGTCTCCCGCTCACAGCCCTCG 4680  
 QY 4681 GCGTCCCTCCCATTCAGAGTGGCTGTCTCATAGAGCAGCAGCAACATGAGTGGGAGCAT 4740  
 Db 4681 GCGTCCCTCCCATTCAGAGTGGCTGTCTCATAGAGCAGCAGCAACATGAGTGGGAGCAT 4740  
 QY 4741 GAGTCTTTGCTGACCTCCCTAGCGGAGAGCAATTTCTTGAGTTAGATAGG 4800  
 Db 4741 GAGTCTTTGCTGACCTCCCTAGCGGAGAGCAATTTCTTGAGTTAGATAGG 4800  
 QY 4801 TTTAGGAAAAAAGACATGCAACCAACGGAATCAATAGAGTTGAAGATTAAGTAC 4860  
 Db 4801 TTTAGGAAAAAAGACATGCAACCAACGGAATCAATAGAGTTGAAGATTAAGTAC 4860  
 QY 4861 TTTGTTTGGCTGTTTTTTTATGTTCTATATGATTTTGAAGATTAAGTAC 4920  
 Db 4861 TTTGTTTGGCTGTTTTTTTATGTTCTATATGATTTTGAAGATTAAGTAC 4920  
 QY 4921 AGATGTTGTTCAATTAATTCCTTTTATGAAAAAGACCTGTAAGATTAAGTAC 4980  
 Db 4921 AGATGTTGTTCAATTAATTCCTTTTATGAAAAAGACCTGTAAGATTAAGTAC 4980  
 QY 4981 TAGGAGAAACTGTCTTCATTCATTTGATGATTTGATCAATTTGATCACTCAACCACT 5040  
 Db 4981 TAGGAGAAACTGTCTTCATTCATTTGATGATTTGATCAATTTGATCACTCAACCACT 5040  
 QY 5041 CTTTTTAAAAA 5055  
 Db 5041 CTTTTTAAAAA 5055

RESULT 2  
 HSA250408  
 LOCUS HSA250408 4743 bp DNA linear PRI 21-OCT-1999  
 DEFINITION Homo sapiens GLI3 gene for GLI3 protein.  
 ACCESSION AJ250408  
 VERSION AJ250408.1 GI:6102811  
 KEYWORDS GLI3 gene; GLI3 protein; transcription factor.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 4743)  
 AUTHORS Kalif-Suske M., Wild A., Topp J., Messling M., Jacobsen E.M., Bornholdt D., Engel H., Heuer H., Kalls C.M., Ausems M.G., Barone R., Herzog A., Heutink P., Homfray T., Gilleßen-Kaesbach G., Konig R., Kunze J., Meinecke P., Muller D., Rizzo R., Streng S., Superl-Furze A. and Grzeschik K.H.  
 TITLE Point mutations throughout the GLI3 gene cause Greig cephalopolysyndactyly syndrome  
 JOURNAL Hum. Mol. Genet. 8 (9), 1769-1777 (1999)  
 MEDLINE 99371786  
 PUBMED 10441342  
 REFERENCE 2 (bases 1 to 4743)  
 AUTHORS Kalif-Suske M.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-OCT-1999) Kalif-Suske M., Zentrum fuer Humangenetik, Philipps-Universitaet Marburg, Bahnhofstrasse 7a, 35037 Marburg, GERMANY  
 COMMENT Related sequences M57609 and M34366.  
 FEATURES  
 source  
 1..4743  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_type="lymphocytes"  
 1..4743  
 /gene="GLI3"  
 1..4743  
 /gene="GLI3"  
 /function="Transcription factor"

```

/codon_start=1
/product="GLI3 protein"
/db_xref="GI:6102812"
/translation="MEAQSHSTTEKKKVENSIKSTRDVSSEKAVASSTTSNEDE
SQGTVHRERRNAITQPONVGLSKVSEESTSSDERASLIKKEHSLPHVAPSV
PYRGTFVMDPRNGMEPHYHPHLPFAPHPVPIDARHHEGRYHYDPSPIPLHMTS
ALSSPTLPOLPFIIRSPHNPAAASESPSPHPYINPYMDYIRSLSHSPSLMISA
TRGLSPDAPHAGVPAEYHOMALLTGORSYADIIPSAATAGTCAIHEYLHMDH
TRFSSPLSARSFKRTLISPLSDHSDPLOTMTISNSLVTIILNRRSSSSAGSY
GHLASAIAPALSTISSAPVSLUMHQQLLSQQLSGAFHSPLIHPAPTFPTQRP
IGPITVMPVNVSGPSSSEKQPTSEAVSTGDMHNRKSLIKPDELPSPGARG
QOEQPTVTLVREEDKDESKQPEVIYETNHCWECCAREFTQQLVHHIINNDIHG
EKKEFYKWLDCSRQKPPKAOYMLVHMRRITGEKPKHCTEGCKTKYRLENLKH
LRSHTKEXYVCEHEGKNKAFSNADRAKHONRTHSNEKPYVCKIPGCTKYRTPDSSL
RKHVATVHGPAAHVKKORGDJHPPRPPRSDGSHSQSRSPGPTQALGEOQDLSNT
TSKRECLQVTKVKAEPMTSPSPGQSGSSQSPISNYNSGLEMLPLTDGSGIG
LSAIDETPTMDSTIATATLALQARNPAGTKMEHVKLERLKVONMFPRLNPILP
PKAPAVSLIGNTOSNNCTSLGGMTLLPGRSDLSGVDMVLMNLRSDSASTISS
AYLSRRSSGIPCFSSRRSSSEASQAEGRPQNSVADSDPISTDSARRSSEASQDG
LPSLLSLTPAQYRLKAKYAAATGGPPPTPLPNERMSLTKRLALLGDALPEGVALPP
VHAPRCSDDGAGHYGRHLOPHDALGHGVRASDPVRTGSEGLALPVPRPSSLSC
NPPMATSAEKSLSLVLONTREPGGSRNFPSSPPSITENVTLDETMDADANLD
EDFLPDVVYVINSQNGAGEHFSALPDDSKVPHGPDGEPDAPGLPDSHAGQQPHAL
EQPCPEGSTDLPIQWNEVSSGADLSKLCGRPAVPQPTRAFPCNGVMVHPQNP
LRSGPAGGYQTIGENSNPYGPEHLMLHNSPGSGTSGNAFHEOPCKAPQYGNCLNRP
VAPGALDGACGAGIOASLKSTPMOQSGGOLNFGPLVPAPNESAGSMVNGMONDPVGO
GYLAHOLLGDSWQHPGACGPRGQMLGOISATSHINIYQPECLPAGHMGQSPLSLA
VVRGQYQSPGSGRRQAMPRLSALQSGQSDTSQTCRVNKGIMKEMGQVHTVDSL
QNYSGOFTDVTGFSQODTKAGFSISDASCLLOGTSAKSELLSPGANQVHTVDSL
DHDHLEQVIDFDADIIDGDHSLSLMSLALSPSIIQNLSSSSRLTTPRASLPFPALSM
SITNMAIGDMSLLTSLADESKFLAVMQ"
BASE COUNT      1141 a 1540 c 1233 g 829 t
ORIGIN

```

```

Query Match      93.6%; Score 4732; DB 9; Length 4743;
Best local similarity 100.0%; Pred. No. 0;
Matches 4743; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 55 ATGGAGGCCAGTCCACAGCTCCAGCACCACCTGAAAGAAAAGATTGAGAAATCCATA 114
DB 1 ATGGAGGCCAGTCCACAGCTCCAGCACCACCTGAAAGAAAAGATTGAGAAATCCATA 60
QY 115 GTGAAGTGTCTCCACTCGAACAGATGTGAGCGAGAAAGCGCTTGCCTCCAGCACCACTTCT 174
DB 61 GTGAAGTGTCTCCACTCGAACAGATGTGAGCGAGAAAGCGCTTGCCTCCAGCACCACTTCT 120
QY 175 AATGAGGATGAAAGTCTTGACAGACTTATCACAGAGAGAGAGAAACGCAATCATTATG 234
DB 121 AATGAGGATGAAAGTCTTGACAGACTTATCACAGAGAGAGAGAAACGCAATCATTATG 180
QY 235 CAGCCACAGAAATCTCAGGGGCTCAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 294
DB 181 CAGCCACAGAAATCTCAGGGGCTCAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240
QY 295 GAGAGGCCCTCATGTATCAAGAAAGAGATCCATGGGTCCCTGCCACACCTGCGGAGGCC 354
DB 241 GAGAGGCCCTCATGTATCAAGAAAGAGATCCATGGGTCCCTGCCACACCTGCGGAGGCC 300
QY 355 TCTGTGCCGTACCGGGAGCGGTGTTGGCATGGAGCCCGCAGGAATGGTTACATGGAGCCC 414
DB 301 TCTGTGCCGTACCGGGAGCGGTGTTGGCATGGAGCCCGCAGGAATGGTTACATGGAGCCC 360
QY 415 CACTACACCCCTCTCATCTTTTCCCTGCTTCCATCTCCCTGATACCAATTCATGTCAGAC 474
DB 361 CACTACACCCCTCTCATCTTTTCCCTGCTTCCATCTCCCTGATACCAATTCATGTCAGAC 420
QY 475 CATCATGAGGCGGCTTACCATTCAGATCCATCTCCGATTCCTCCATTCGATGATGACTTCC 534
DB 421 CATCATGAGGCGGCTTACCATTCAGATTCATCTCCGATTCCTCCATTCGATGATGACTTCC 480
QY 535 GCCTTATCTAGTAGCCCTTAGTATCCGAGCTCCGCTTCATTAGGATCTCCCCACACCGG 594
DB GCCTTATCTAGTAGCCCTTAGTATCCGAGCTCCGCTTCATTAGGATCTCCCCACACCGG 1620

```



QY 1675 ACGGGGAGAACGCTCACAATTTGCACTTTTGAAGTTGACAAAGGCTACTTCGAGACTA 1734  
Db 1621 ACGGGGAGAACGCTCACAATTTGCACTTTTGAAGTTGACAAAGGCTACTTCGAGACTA 1680  
QY 1735 GAAACTTGAACACACACTTGAGATCTCACACTGAGAGAAACATAGCTGTGTAGACAC 1794  
Db 1681 GAAACTTGAACACACACTTGAGATCTCACACTGAGAGAAACATAGCTGTGTAGACAC 1740  
QY 1795 GAAGTTGCAACAAGGCTTTCCTCAATGCTCTGATCGCGCCCAACACACAAACAGAACG 1854  
Db 1741 GAAGTTGCAACAAGGCTTTCCTCAATGCTCTGATCGCGCCCAACACACAAACAGAACG 1800  
QY 1855 CATTCOAATGAGAAACCATATGTGTGCAAAATCCAGAGCTGCACCTAAGCGTTTACACAG 1914  
Db 1801 CATTCOAATGAGAAACCATATGTGTGCAAAATCCAGAGCTGCACCTAAGCGTTTACACAG 1860  
QY 1915 CCAAGTCCCTCGGAAACATGTGAAGACAGATGCAATGCGCCAGAGGCTCATGTACCAAG 1974  
Db 1861 CCAAGTCCCTCGGAAACATGTGAAGACAGATGCAATGCGCCAGAGGCTCATGTACCAAG 1920  
QY 1975 AAGCAGGAGGAGACATCCATCCTCGCGCCGACCCCGAGAGATTCGCGCAGCCATTCA 2034  
Db 1921 AAGCAGGAGGAGACATCCATCCTCGCGCCGACCCCGAGAGATTCGCGCAGCCATTCA 1980  
QY 2035 CAGTCCAGGCTGCGCGACCGACTCAGGAGAGCCCTTGTGAGCAGCAGGACCTCAGC 2094  
Db 1981 CAGTCCAGGCTGCGCGACCGACTCAGGAGAGCCCTTGTGAGCAGCAGGACCTCAGC 2040  
QY 2095 AACACTACCTCAAGGGGGAAGAAATGCTTCAGAGTGGAANAACCGTCAAGCGAGAGACCA 2154  
Db 2041 AACACTACCTCAAGGGGGAAGAAATGCTTCAGAGTGGAANAACCGTCAAGCGAGAGACCA 2100  
QY 2155 ATGACATCTCAGCAAGCCCTGCTGCTCAGTCTTCATGACAGCAGCAACAGTCCCGCCATC 2214  
Db 2101 ATGACATCTCAGCAAGCCCTGCTGCTCAGTCTTCATGACAGCAGCAACAGTCCCGCCATC 2160  
QY 2215 AGCAACTATTCAACAGTGGGCTCGAGCTTCTCTGACCGATGGAGTAGTATAGGAGAC 2274  
Db 2161 AGCAACTATTCAACAGTGGGCTCGAGCTTCTCTGACCGATGGAGTAGTATAGGAGAC 2220  
QY 2275 CTCAGTGCCATCGATGAACCCCAATCATGAGTCAACCATTTCCACTGCAACACAGCC 2334  
Db 2221 CTCAGTGCCATCGATGAACCCCAATCATGAGTCAACCATTTCCACTGCAACACAGCC 2280  
QY 2335 CTGTCTTTGCAAGCCAGGAGAAACCGCGCAGGACCAATGATGAGCAGCAGTAAACTA 2394  
Db 2281 CTGTCTTTGCAAGCCAGGAGAAACCGCGCAGGACCAATGATGAGCAGCAGTAAACTA 2340  
QY 2395 GAAAGCTTAAACAAAGTGAATGGAATGTTTCCGCGACTGAAACCCCATTTACCCCTAAA 2454  
Db 2341 GAAAGCTTAAACAAAGTGAATGGAATGTTTCCGCGACTGAAACCCCATTTACCCCTAAA 2400  
QY 2455 GCCCTGCGGTCTCTCTCATAGGAATGGACACAGTCCAAACACACTTCGAGCTTG 2514  
Db 2401 GCCCTGCGGTCTCTCTCATAGGAATGGACACAGTCCAAACACACTTCGAGCTTG 2460  
QY 2515 GTGGGCCCATGACGCTTCTCCGGGCAAGACGACCTCTCTGGGTGAGCTCACTATG 2574  
Db 2461 GTGGGCCCATGACGCTTCTCCGGGCAAGACGACCTCTCTGGGTGAGCTCACTATG 2520  
QY 2575 CTGAACATGCTCAACAGAGGAGACAGCAGCGCCAGCACCATCAGCTCGCGCTTACCTGAGC 2634  
Db 2521 CTGAACATGCTCAACAGAGGAGACAGCAGCGCCAGCACCATCAGCTCGCGCTTACCTGAGC 2580  
QY 2635 AGCCGCGCTCTCAGGATCTCGCCCTGCTTCTCCAGCGCCGCTCCAGCAGGCGTCA 2694  
Db 2581 AGCCGCGCTCTCAGGATCTCGCCCTGCTTCTCCAGCGCCGCTCCAGCAGGCGTCA 2640  
QY 2695 CAGGCGAGGCGCGCGCAGACGAGTGGCGGAGCTCTCTACGACCCCATCTCCACC 2754  
Db 2641 CAGGCGAGGCGCGCGCAGACGAGTGGCGGAGCTCTCTACGACCCCATCTCCACC 2700

QY 2755 GACCCCTCGCGCCCTCCAGGAAGCCAGCCAGACGCGCCTGCCAGCCTGCTCAGC 2814  
Db 2701 GACCCCTCGCGCCCTCCAGGAAGCCAGCCAGACGCGCCTGCCAGCCTGCTCAGC 2760  
QY 2815 CTCACGCCGCCAGCAGTACCGCTCAAGGCCAAGTACGCGCTGCCAGGAGGCGC 2874  
Db 2761 CTCACGCCGCCAGCAGTACCGCTCAAGGCCAAGTACGCGCTGCCAGGAGGCGC 2820  
QY 2875 CCGCGCAGCGCCCTGCCCAACATGAGAGGATGAGCCTGAAGCAGCGCCTGGCGCTC 2934  
Db 2821 CCGCGCAGCGCCCTGCCCAACATGAGAGGATGAGCCTGAAGCAGCGCCTGGCGCTC 2880  
QY 2935 GGGATGCTCGAGCCTGCGGTGGCCCTGCTCCAGTTCATGCCCCGAGAGGTGCAG 2994  
Db 2881 GGGATGCTCGAGCCTGCGGTGGCCCTGCTCCAGTTCATGCCCCGAGAGGTGCAG 2940  
QY 2995 GACGGGAGGCCACCGCTACGGGCGCGCCACCTGACGCGCAGCATGCGTGGGCCAC 3054  
Db 2941 GACGGGAGGCCACCGCTACGGGCGCGCCACCTGACGCGCAGCATGCGTGGGCCAC 3000  
QY 3055 GGCCTGAGGAGGCCACCGCTGCGGACAGGCTCCGAGGCGCTGGCCCTGCTCGT 3114  
Db 3001 GGCCTGAGGAGGCCACCGCTGCGGACAGGCTCCGAGGCGCTGGCCCTGCTCGT 3060  
QY 3115 GTGCCGCTTCAGCAGCTCAGCAGTGCACACCCCGCGGATGCGCCAGTCCGCGGAG 3174  
Db 3061 GTGCCGCTTCAGCAGCTCAGCAGTGCACACCCCGCGGATGCGCCAGTCCGCGGAG 3120  
QY 3175 AAGCGCAGTCTCGTCTTCAGATTACACGCGCGCCGAGGCGCCAGTCCCGAACTTC 3234  
Db 3121 AAGCGCAGTCTCGTCTTCAGATTACACGCGCGCCGAGGCGCCAGTCCCGAACTTC 3180  
QY 3235 CACTCGTCCCGCTCTCTCCAGCATCACGGAACCTTCACCTGGAGTCCCTGACCATG 3294  
Db 3181 CACTCGTCCCGCTCTCTCCAGCATCACGGAACCTTCACCTGGAGTCCCTGACCATG 3240  
QY 3295 GACGCTGATGCCAACTTCAGAGTGAAGATTTCTGCGGACGAGCTGGTGCAGTATTTA 3354  
Db 3241 GACGCTGATGCCAACTTCAGAGTGAAGATTTCTGCGGACGAGCTGGTGCAGTATTTA 3300  
QY 3355 AATTCCAGAACCAAGCAGGTACGAGCAGCACHTTCCAGCGCCCTCCCGAGCAGCAGC 3414  
Db 3301 AATTCCAGAACCAAGCAGGTACGAGCAGCACHTTCCAGCGCCCTCCCGAGCAGCAGC 3360  
QY 3415 AAAGTGCCCGCCGCGCTGACTTTGACGCGCCCGGCTGCCAGACACGCGCTGGC 3474  
Db 3361 AAAGTGCCCGCCGCGCTGACTTTGACGCGCCCGGCTGCCAGACACGCGCTGGC 3420  
QY 3475 CAGCAGTTCCATGCTCGAGCAGCCCTGCCCGAGGCGAGCAAAACCGACCTGCCCAT 3534  
Db 3421 CAGCAGTTCCATGCTCGAGCAGCCCTGCCCGAGGCGAGCAAAACCGACCTGCCCAT 3480  
QY 3535 CAGTGGAAAGCAGTCACTCCGGAAGCGCGACCTGCTCTCTCAAGCTCAAGTGGG 3594  
Db 3481 CAGTGGAAAGCAGTCACTCCGGAAGCGCGACCTGCTCTCTCAAGCTCAAGTGGG 3540  
QY 3595 CCGCGGCCGCTGTGCGCAGACTCGCGCCCTTTGGTTCCTGCAACGCACTGGTGCAC 3654  
Db 3541 CCGCGGCCGCTGTGCGCAGACTCGCGCCCTTTGGTTCCTGCAACGCACTGGTGCAC 3600  
QY 3655 CCGCAGAACCCCTTTCAGAGGCGGCTGCTGGGGGCTATCAGACCTTCGGGAGAACAGC 3714  
Db 3601 CCGCAGAACCCCTTTCAGAGGCGGCTGCTGGGGGCTATCAGACCTTCGGGAGAACAGC 3660  
QY 3715 AACCCCTACGCTGGCCAGACACTTGATGCTCCACAAACAGCCCGGAGTGGCAGCAGT 3774  
Db 3661 AACCCCTACGCTGGCCAGACACTTGATGCTCCACAAACAGCCCGGAGTGGCAGCAGT 3720  
QY 3775 GGAACGCTTTCATGAACAGCCCTGTAAAGCCCGCAGTATGGAACTGTCTCAACAGG 3834  
Db 3721 GGAACGCTTTCATGAACAGCCCTGTAAAGCCCGCAGTATGGAACTGTCTCAACAGG 3780  
QY 3835 CAGCCAGTGGCCCTTGTGCTGCTGACGCGGTGCTGTCGCCGGATTCAGACCTCAAG 3894



```

|||||
Db 3781 CAGCCAGTGGCCCTGGTCACTCGACGCTGCTGTGGTGGCGGATTCAGACCTCAAG 3840
QY 3895 CTGAACAGACACCCCATCATCAAGAGGAGCGGGGCCAGCTGAATTTCCGGCTCGCGTAGG 3954
Db 3841 CTGAAGAGACACCCCATCATCAAGAGGAGCGGGGCCAGCTGAATTTCCGGCTCGCGTAGG 3900
QY 3955 CCAATAGTCACTGGCAGCAGTGGTGAATGCATGCAGAACAGACAGCCAGTGGACAG 4014
Db 3901 CCAATAGTCACTGGCAGCAGTGGTGAATGCATGCAGAACAGACAGCCAGTGGACAG 3960
QY 4015 GGTACCTGGCTACACAGCTCTCGGCGACAGCATGCAGACACCGGGGCGAGCCGCCGCC 4074
Db 3961 GGTACCTGGCTACACAGCTCTCGGCGACAGCATGCAGACACCGGGGCGAGCCGCCGCC 4020
QY 4075 GGTACAGAGTCTGGGAGATAGTGTACCTACACATCAACATCAACATCAACAGGGCCA 4134
Db 4021 GGTACAGAGTCTGGGAGATAGTGTACCTACACATCAACATCAACATCAACAGGGCCA 4080
QY 4135 GAGAGTCTGGCTGACAGGGCTCAGGCATGGGCGAGCCAGCCGCTCAAGCTTGGCAGTTGTC 4194
Db 4081 GAGAGTCTGGCTGACAGGGCTCAGGCATGGGCGAGCCAGCCGCTCAAGCTTGGCAGTTGTC 4140
QY 4195 AGGGGTACACAGCATGTGCCAGCTTTGGGGCAGCAGCGCCAGGCTATGCCGAGGGAC 4254
Db 4141 AGGGGTACACAGCATGTGCCAGCTTTGGGGCAGCAGCGCCAGGCTATGCCGAGGGAC 4200
QY 4255 AGCTTGTCTGAGTCAGGACAGCTCAGTGACAGCAAGTCAGACCTGCAGGCTGAATGTT 4314
Db 4201 AGCTTGTCTGAGTCAGGACAGCTCAGTGACAGCAAGTCAGACCTGCAGGCTGAATGTT 4260
QY 4315 ATCAAGATGAGATGAAGAGGGCAGCCCATCGCTGTCTAATCTGCAGAAATTAATCTCT 4374
Db 4261 ATCAAGATGAGATGAAGAGGGCAGCCCATCGCTGTCTAATCTGCAGAAATTAATCTCT 4320
QY 4375 GGTCACTTATGACCAAAACCGTGGCTTCAGTCAGCAAGACAGCAAGACCTGGTTCATTC 4434
Db 4321 GGTCACTTATGACCAAAACCGTGGCTTCAGTCAGCAAGACAGCAAGACCTGGTTCATTC 4380
QY 4435 TCATTATTCAGACCCAGCTGCTGTCTACAGGGGACAGCGCCCAAAACCTCTGAGTTACTT 4494
Db 4381 TCATTATTCAGACCCAGCTGCTGTCTACAGGGGACAGCGCCCAAAACCTCTGAGTTACTT 4440
QY 4495 TCCCGAGGTCTAATCAGCTGACCAACAGCAGTGGACAGCTCGACAGCCATGACCTGGAA 4554
Db 4441 TCCCGAGGTCTAATCAGCTGACCAACAGCAGTGGACAGCTCGACAGCCATGACCTGGAA 4500
QY 4555 GGGGTACAGATGACTTCGATGCCATCATAGAGATGGGACACATCCAGCCTGATGTCG 4614
Db 4501 GGGGTACAGATGACTTCGATGCCATCATAGAGATGGGACACATCCAGCCTGATGTCG 4560
QY 4615 GGGGCTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGCTCACCAG 4674
Db 4561 GGGGCTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGCTCACCAG 4620
QY 4675 CCTCGGCGCTCCCTCCATTCACCTCCAGTCGCTGTCCATGAGCACCACCAACATGCTATCGG 4734
Db 4621 CCTCGGCGCTCCCTCCATTCACCTCCAGTCGCTGTCCATGAGCACCACCAACATGCTATCGG 4679
QY 4735 GGACATGAGTTCTTGTGCTACCTCCCTAGCGGAAGAACAAATTCCTTGCAGTTATGCA 4794
Db 4680 GGACATGAGTTCTTGTGCTACCTCCCTAGCGGAAGAACAAATTCCTTGCAGTTATGCA 4739
QY 4795 ATAG 4798
Db 4740 ATAG 4743

```

RESULT 3  
 MMGLI3GEN  
 LOCUS  
 DEFINITION M.musculus mRNA for Gli3 protein.  
 ACCESSION X95255

```

X95255.1 GI:1419013
GLI3 gene; Gli3 protein.
Mus musculus.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Thien, H., Buscher, D. and Ruther, U.
Cloning and sequence analysis of the murine Gli3 cDNA
Biochim. Biophys. Acta 1307 (3), 267-269 (1996)
9605347
PUBMED 868459
REFERENCE 2 (bases 1 to 5113)
AUTHORS Thien, H.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-1996) H. Thien, Medizinische Hochschule Hannover,
Institut fuer Molekularbiologie, OE 5250, Konstanty-Gutschow-Str.
8, D-30
Location/Qualifiers
FEATURES
source
1..5113
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="13"
172..4962
/gene="Gli3"
172..4962
/gene="Gli3"
/codon_start=1
/product="Gli3 protein"
/protein_id="CAA64543.1"
/db_xref="GI:1419014"
/db_xref="MGD:MGI:95729"
/db_xref="SWISS-PROT:O61602"
/translation="MEQAHSSTATERRKKAENSGIKCPTRTFDVSEKAVASSTTISNEDE
SPGOIYHRRRNATMQPVSQVQLGKISSEPFSSSSDERASLIKEHGLHSPLHAEPSL
PYRTGFAMPDRNGMYPHPLFPFAPFPVPIDARHHEGYHYDPSPIPLHVPSP
ALSSPTYPDLPIFIRISPHRNPTAAESPFSPHPYINPYNDYIRSLHCSPLSMISA
ARGSLPTADPHAGVSPAEYHOMALLTGQSPYADILPSAATAGAGAHIMHEYLHAMD
TRFSPRLSPASPSRKRLTSLPSLDHSDFDLQTLRTSPNSLVTILNRSNSASGSGY
GHLASAIAPALSFTYPSAPVSLHMQOILSRQSLGSAFGHSPPLIHPAPTPTOPR
IPGLPTVLPVQVSSGSESSQKPTSESASVSTGGPMHNRKRIKIDEDLPFGSRG
QQEQQETTLVKEADKDESKQEPVIYETNCHWEGCTREFDQDLVHLINDDHIG
EKKEFVCRWLDCSEHQKFAQYLVVHRRHTGEKPKCTFEGCTKAYSRLENLKH
LRSHGTGKPYVCEHEGNCNKAFLSNADRAKHQNRTHSNEKPYVCKIPCTCAFGAEKELSN
RKHYTVHGEPAHYTKKQGMHPHPPPPRDSGSHSOSRSPGPRGTCAGAEKELSN
TSKPEECLOVKTVAEKDPTSQSPGSSQSSQSPISNYSNSELRLPLTDGGSVAD
LSAIDEPTMDSTISTATALLAOLARRNPAGTKWEHIKLERLKQVNGMPTRLNPILP
SKAPAVSPLINGNGTQSNNNYSSGGPTLLPSPDLSGVDFTVLNTLNRRDSNTSTISS
AYLSRRSSGISPCFSRRSSEASQAEGRPNQVSVADSYDPISTDASRRSSEASQDGG
LPSSLSTPVQOYALKAKYAPTGGPTPLPHMERLSLTKMALLGEGRDQSGVTLP
VHPPRCSGGGHTYRGRHLMPHDALANSVRRSDPVRTVSENMSLARVORFSLNSF
NPNLPSPVEKRSILVQNTYRTOESSOPRYFOACPSPITENVALEALTDADANLND
EDLLPDDVQVYLNQNTGYCQQLQSGISESKVAHEPEDLDLAGLPDHSVQGEYAL
EQPCSEGGKTDLPLOWNEVSSGTDLSLKKCCQRPQRPQRPQGFGLYNNVHHPHL
WKVGTGAGGYQTLGENTSYNGPEHFAIHSGDGLGTNGNTHFEQPKTQOYQSQLNR
QPLFSSALDHACGTGIGSKLNGSLNGGLDLSLVAPNELAGTNGVMTQDMQ
GOGYIAHOLLGSMQHGSPRPGQVVGQVATSHINIYOTSCSLPCTQDNSSQPS
MAALRGYQPCASYGNNRQAMPNRLTLQOGLSDMSQSRVNSIKMEAGQSQQLCS
TVQNSQGFYDQTMFGQQQDKKAGSFLSDANCLLQGTCTENSELLSPGANQVTSYVD
SFESHDLLEGVQIDFADLDIDGHTSLMSGALSPSIIQNLHSHSRSLTTPRASLPFP
IHGHQHGHYGVEFFADLPCKRRKQVPCSYAVGGRQGGPQTRLK"
BASE COUNT 1337 a 1490 c 1254 g 1032 t
ORIGIN

```

Query Match 73.1%; Score 3696.6; DB 10; Length 5113;  
 Best Local Similarity 84.2%; Pred. No. 0;  
 Matches 4212; Conservative 0; Mismatches 769; Indels 24; Gaps 3;  
 QY 9 CGTGGCATTTTGTCTCGAAGAGACTGAGCTAAATGAGAAGACATCATGAGGCCCTCAGT 68  
 Db 126 CCAGGGGATTCCTTTGAGAAACAGCTGAAGTAATGAGAAGACATTTATGGAGGCCAGCC 185  
 QY 69 CCACAGCTCCAGCAGCAGCTGAAAGAAAAAGTTGAGAATTCATAGTAGTGAAGTGTCCAC 128





```
Db 4566 AGATGCCAACTGCGCTGCCAAGGACATGACTGAAACTCTGAGTTACTCTCCCGAGG 4625
Qy 4503 TGCTAAATCAGGTGACAGCAGCAGCTCGACAGCCATGACCTGGAAGGGTACA 4562
Db 4626 TGCTAACCGAGTACACAGCAGTGTGACAGCTTTGAGAGTCATGACCTAGAAGGTGCA 4685
Qy 4563 GATTGACCTCGATGCCATCATAGACAGTGGGACCACTCCAGCCTGATGTCGGGGGCCCT 4622
Db 4686 GATTGATTTTATGATGCCATCATAGATGATGGGACCATACAGCCTAATGTCAGGGCCCT 4745
Qy 4623 GAGCCCAAGTATCATTCAGAACCTTTCCCATAGTCTCCCGCCTCACAGCCTCGGGC 4682
Db 4746 GAGCCCAAGTATCATTCAGAACCTTTCCCATAGTCTCCCGCCTCACAGCCTCGGGC 4805
Qy 4683 GTCCCTCCCATTCCTCCAGTCTGCTCCATGACACCAACCAATGGCTATCGGGGACATGA 4742
Db 4806 ATCCCTCCCATTCCTCCATGCTCCATGACACCAACCAATGGCTATCGGGGACATGA 4865
Qy 4743 GTCTTTGCTGACCTCCATAGCAGGAAAGAAATTCCTTGCAGTTATGCAATAGGCTT 4802
Db 4866 GTCTTTGCTGACCTCCATGCTCCATGACAGAAAGAAATTCCTTGCAGTTATGCAATAGGCTT 4925
Qy 4803 TAGGAAAAGAGCTCAACCAACGAGAAATCAATAGAGTTGAAGATTAACACTGACTT 4862
Db 4926 TAGCAAGGAGGACCAACCAACGAGAAATCAATAGAGTTGAAGATTAACACTGACTT 4969
Qy 4863 TGTTTGGCTGTTTCTTGTAGTCTGATGATTTTGTAGCAATCTCATCTCACCTAACTAG 4922
Db 4970 -TTTTTTTCTTCTTCTTGTAGTCTGATGATTTTGTAGCAATCTCATCTCACCTAGTGG 5028
Qy 4923 ATGTGTTCAATATATATCTTTTATGGAAGAGCTCTGAAACCCCTAAAGTATCTTA 4982
Db 5029 ATGTGTTCAATATATATCTTTTATGGAAGAGCTCTGAAACCCCTAAAGTATCTTA 5088
Qy 4983 -GGGAGAAACTGCTTCCATTTTCAG 5006
Db 5089 GGGAGAAACTGCTTCCATTTTCAG 5113

FEATURES
     CDS             4982 bp      mRNA      linear      VRT 18-JUN-2000
     DEFINITION      Coturnix coturnix kruppel family member GLI3 (Gli3) mRNA, complete
     ACCESSION       AF231112
     VERSION          AF231112.1
     KEYWORDS         Coturnix coturnix.
     ORGANISM         Coturnix coturnix.
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
     Phasianinae; Coturnix.
     1 (bases 1 to 4962)
     Borycki, A., Brown, A.M. and Emerson, C.P. Jr.
     Shh and Wnt signaling pathways converge to control Gli gene
     activation in avian somites
     Development 127 (10), 2075-2087 (2000)
     20233829
     MEDLINE          10769232
     REFERENCE        2 (bases 1 to 4962)
     Borycki, A., Brown, A.M. and Emerson, C.P.
     Direct Submission
     Submitted (03-FEB-2000) Cell and Developmental Biology, University
     of Pennsylvania, 421 Curie Blvd, Philadelphia, PA 19104-6058, USA
     COTURNAL
     LOCATION/Qualifiers
     1. .4962
     /organism="Coturnix coturnix"
     /db_xref="taxon:9091"
     /dev_stage="stage HH12 embryo"
     1. .4962
     /gene="Gli3"
     200. .4933

     gene
     CDS
```

```
/gene="Gli3"
/Note="zinc finger transcription factor"
/codon_start=1
/product="Kruppel family member GLI3"
/protein_id="AAF76851.1"
/db_xref="GI:8571370"
/translation="MEAQSHSSTTTTEKKKVENSVKCSRTDVSSEKAVASSTTTSNEDP
SPGOTYHRRERNAITMPOGGGGLSKISEPSTSEERASLIKKEIHGSHLPEPSV
PYRGLTMDPRNGYMDPHYRDPHLPFAFPHPVIDARHHEGRYHHEPSPPIPLHPV
AALSSPTISDLPIRISPHRNAASASESPSTPHYPINPYMDYIRSLHSSPSLSMIS
SARSLSPTDAPHAGVPAEYHYOMALLAGORSYADIIPSAATAGALAHMEYLHAMD
YHLSASAISSARPSRKRTLSISPLSDHSDFLQITWRTSPNSLGLFLLNNSRSSASG
YHLSASAISSALSFYTPPTVPSLQOHHQIISROOTLGSAGFSPPLIHPAPFTPQ
RPIPIPSVLNPVOVSSGSTOQNEKPTSESAVSTGDPHMKRSKIKPEDLPSG
AGVQEOPEGMTLYKRGDDKQEPVEYETNCHWEGSKREFTDQRLVHLNNH
IHGSKFEVCRWLDLSREKPFKAQYMLVHMRRHTGKPHKCTFECTKAGYSLRNLH
KTLHRSHTGEKPYVCEHEGKNKAFNSDRAKHQNRTHSHNEKPYVICTKRYTDP
SLRKHVKTVHGPEAHVTKKQGDIIHPPPPPDPGSHSQTSPGQOCTGATGQKDL
NNTLSRDEECPIMDSITATTAALQARRMTGTKMEQVKLERLQKVGMLPRLNP
VGLSVIDEPTIMDSITATTAALQARRMTGTKMEQVKLERLQKVGMLPRLNP
VPSKAPTLPPLIGNGTKSNSCHVGSMTLLPNRNLSTDTITVNLNRRSNTST
ISSAYLSRRSSGSPCFSSRRSDASQAEGRPNVSADSYDPTSDARRSEASQ
CDDPLSLSTPAQYRLKAKYAAATGGPPTPLPNMRSLKTMALLDGDRESGV
PLPVPNAPRRSSDGGANGYRRHLLSHDALNGTTRASDPVPMVSDSQSPVQHFNS
LNSPSPALPPSMKRNRLVQNTSRSEGLFRGFSPPSPISINVTLEADVAEAGS
LNDEDLPLDDVQYLNQONQMYDHLNLDNKNQHSVVLGNPNPSSFDRAVPASS
OPAGSEASKDLPIONNEVSSGSELSPTKLKCSORSAGOOARFELYNMVOOONL
ORSNVQONQYOSILMENSQYVILQONTGLNGASSFSMPNPKPYSESISROPMSA
MDNSCVTVQVQKLRSSNMPASGQNPQHPITSSGGQATDMQNRNMGGQYILN
QLIGDGLHYGVNQSGPMGLQVPSYSLYQGPQSPVFTHTGNSGLSVAKSY
QPCANYSGNRQNRVLRNLTQOQGHVSDYNOTCRVNTIKMEMQGSQOQFCNNQYSG
QLYDPMGFSHQAMKTSFVSEANLLQETATANSSELLSPGANGSVSTVDSLDNN
SLEGOVIDFAIIDDGDHVSILSGALSIIQNLRSRSLTTPRASLTTPAVPVSTT
NMAITGDMSSLLTSLAESKFLAVM"
BASE COUNT      1401 a 1334 c 1175 g 1052 t
ORIGIN
Query Match      54.6%; Score 2758.6; DB 5; Length 4962;
Best Local Similarity 75.1%; Pred. No. 0;
Matches 3627; Conservative 0; Mismatches 1154; Indels 46; Gaps 13;
Qy 17 TTTTGTGCGAAGAGAGCTGAAGTAATGAGAAGACATCATCGAGGCCAGCTCCACAGCT 76
Db 162 TATTAGTTTCAAGGAGAGCTGAAGTAATGAGAAGTCACTCGAGGCCGAGCTCCCATAG 221
Qy 77 CCACGACCACTGAAAGAAAAGTTGAGAATTCATAGTCAAGTCCACTCCACTGCAACAG 136
Db 222 CTACACACAGAGAGAGAAAGGTGGAGAATTCATTTGAAATGCTTCCAGTCCGACAG 281
Qy 137 ATGTGAGCGAGAAAGCCGCTTCCCTCCAGCACCACTTTCTAATGAGAGTGAAGTCTGGAC 196
Db 282 ATGTGAGTGAAGAGCTTTCCTCCAGCACCACTTCCATGAGAGTGAAGTCTGGAC 341
Qy 197 AGACTTATCAGAGAGAGAGAGAAAGCAATCATATGACGCCACAGATGTCACAGGGC 256
Db 342 AGACTTATCAGAGAGAGAGAGAAAGCAATCATATGACGCCACAGAGTGGGAGGCC 401
Qy 257 TCAGCAAGTCACTGAGGAACCTTCAACATCGAGTGAAGGCCCTCATTTGATCAAGA 316
Db 402 TCAGCAAGTCACTGAGGAGCTTCAACATTCACGAGGAAGAGGCCCTCATTAAGA 461
Qy 317 AAGAGATCCATGGTTCCTGCCACACGCTGGGAGGCCCTCTGTGCCCTACCGGGAGCG 376
Db 462 AAGAGATCCATGGATCTATATCGCACCTTCTCTGAACCTTCTGTACCTTATCGTGGACAT 521
Qy 377 TGTTCCTGCTGACCCCGAGGATGTTACATGAGGCCGCCACTACC---ACCTCTCTATC 433
Db 522 TGTTCCTGCTGACCCCGAGGAAATGTTACATGAGTCTCTATACCCGGGATCCACCTCAC 581
Qy 434 TTTTCCCTGCTTCCATCTCTCTGTACCAATGATGCGACATCATGAGGCGGCTTACC 493
Db 582 TCTTTCAGGCTTCCACCTCTCTGTACCAATGATGCGGACATCATGAGGAGGCTTACC 641
```

QY 494 ATTACGATCCATCGGATTCCTCCATATGATGACTTCCGCTTATCTAGTAGCCCTA 553  
Db 642 ATTATGAACCATCCCAATTCCTCTGATGCTGCTTCTGCTTATCTAGTAGCCCA 701  
QY 554 CGTATCGGACCTCGCCTTCAATAGGATCTCCCAACCGGACCGGCTGCTGCTCGG 613  
Db 702 CATATTGAGATCTCTTTCATATAGAAATTTCCCAACAGCAAAATTCGCTGAGCATCAG 761  
QY 614 AGTCTCCCTTCAGCCCTCCACATCCCTACATTAATCCCTACATGGACTATATCCGCTCT 673  
Db 762 AGTCTCCCTTCAGCATCTCCACCTGATTAATCCCTACATGGACTATATCCGCTCTCC 821  
QY 674 TGCACAGCAGCCATCGCTCTCCATGATCTCAGCAACCGGCTGGCTGAGCCCTACAGATG 733  
Db 822 TGCACAGCAGCCATCCCTGTCATGATCTCAGCAGCCGCTGACTCAGCCCAACAGAGC 881  
QY 734 CGCCCATGAGAGGTGAGCCAGCAGAGATATCATCATGATGGCCCTTGAATGAGGC 793  
Db 882 CTCACATGCTGGGCTCAGTCCAGCTGAATATATCATCAGATGGCTCTGTTGGAGGCC 941  
QY 794 AGCGCAGCCCTATGAGCATATTTCCCTCAGCTGCCACCGCCGACGCGGGCCATCC 853  
Db 942 AGCGCAGCCCTTACGAGCATATTTCCCTCAGCTGCCACAGCTGGAGCTGGTGTCTTC 1001  
QY 854 ACATGGAATATCTTCATGCTATGATAGCAGCAGATTTCTCCAGCCCGCAGGCTGTAGCCA 913  
Db 1002 ATATGGAATATCTTCATGCTATGATAGCAGCAGCAGGTTTCGAGTCCAAATATCAGCTA 1061  
QY 914 GGCCGAGCGAAACGTACACTGCTCATATACACTCTCCGATCATAGCTTTGAGCTTC 973  
Db 1062 GAGCAAGCCGAAAGCGTACGTTCTCTATATCCGCCCTATCTGATCAGAGCTTTGACCTTC 1121  
QY 974 AGACCATGATAAGAGCTCTCCCACTCTTGTGTCAGATTTCTCAATAATTCCTGAGCA 1033  
Db 1122 AGACCATGATAAGAGCTCTCCCACTCTTGTGTCAGATTTCTCAATAATTCCTGAGCA 1181  
QY 1034 GCTCTTCAGCAAGTGGCTCTATGCTCACTTATCTCAAGTGAATTCAGCCCTGCTTGA 1093  
Db 1182 GTTCTCAGCAGCTGGTTCCTATGCGCACTTGTCTCAAGGCGCAATAAGCCCTGCTCA 1241  
QY 1094 GCTTCACCTACTCTTGGGCGCCGCTCTCTC --- CACATGATCAGAGATCTTCAAGC 1150  
Db 1242 GTTTCACATACCTCTCTACACCGGATCTCCCTCAGCAAAATGATCAGCAAAATATAAGTC 1301  
QY 1151 GACAAACAGCTTATAGTTCAGCTTTGGACACAGCCCTCCACATCTCCACCTGCCCA 1210  
Db 1302 GTACGAAACCTTGGGTTCTGCTTTGGACACAGCCCTCTCTCATCTCATCTGCTCA 1361  
QY 1211 CTPTTCCACACAGAGCCCTATTCAGGGATCCCTACGGTCTGTAACCCCGTCCAGGTCA 1270  
Db 1362 CTPTTCTACACAGAGCCCTATCCCTGGCATCCCTAGTGTCTGTAACCCCTGCTCAAGTCA 1421  
QY 1271 GCTCCGCGCTTCTGAGTCTCA --- CAGAACAGCCCGAGTGTGAGTCTGAGTCAAGCA 1327  
Db 1422 GCTCTGAGCTTCTGAGTCTCAGACAGCAAGCAAAACCCAGAGTGTCTGAGTCAAGCA 1481  
QY 1328 GCACTGTGACCGATGCAACAAAGAGTTCAGAGTCAAAACCCGATGAAGACTCTCCCA 1387  
Db 1482 GCACTGGGATCAATGCAACAAAGCTCTCAAGATGAAGCCCGATGAGGACCTCCCA 1541  
QY 1388 GCCCAGGGGCTGGGGSCAGCAGAAACAGCCGAGAAACACCCCTTGAAGAGGAAG 1447  
Db 1542 GCCCGGTGAGGAAGGCTGAGCAACAGCAGCAAGGATGAGCCCTTGAAGAGGAAG 1601  
QY 1448 GGGACAAGATGAAGAAACAGAGGCTGAAGTCAATCTATGAGCAAACTGCCACTGGG 1507  
Db 1602 GGGACAAGATGAAGAAACAGAGGCTGAAGTCTACGAGAGCAACTGCCATTGGG 1661  
QY 1508 AAGGCTCGGAGGAGTTCGACACCAAGAGCAGCTTGTGCAACCATATAAATAAGGACC 1567  
Db 1562 AAGGCTCTCCGGGAGTTTGACACACAGGAACAGCTAGTGTCAATATAAATAAGGACC 1721  
QY 1568 ATATTATGAGAGAGAGGAGTTGCTGTGCAAGTGGTGGTCTGCTCAAGAGAGCAGA 1627

Db 1722 ACATACATGGTGAAGAAAGAGTTTGTGTCGATGCTGACTGTTCCCGGGAGCAGA 1781  
QY 1628 AACCTTTCAAAGCCAGTATATGTTGGTAGTGCATATAGAGACACACAGCGGAGAGC 1687  
Db 1782 AACCTTTCAAAGCCAGTATATGTTGGTAGTGCATATAGAGACACACAGCGGAGAGC 1841  
QY 1688 CTCACAAATGCACTTTTGAAGTTGCACAAAGGCTACTCGAGACTAGAAAACCTTGA 1747  
Db 1842 CACACAAATGCACTTTTGAAGTTGTACAAAGGCTACTCCAGACTAGAAAACCTTGA 1901  
QY 1748 CACACTTGAGATCTCAGACTGAGAGAAACCATACGCTGTGTGAGCAGCAAGGTTGCA 1807  
Db 1902 CGCACTTGAGATCTCAGACTGGGAGAAACCATATGTTGTGAACATGAGGCTGCA 1961  
QY 1808 AGGCTTTCTCAAAATGCTCTGATCGCGCCAAACACCAACAGAGCAATTCCTCAATG 1867  
Db 1962 AAGCTTTCTCAAAATGCTCTGATCGCGCCAAACACCAACAGAGCAATTCCTCAATG 2021  
QY 1868 AACCATATGTTGCAAAATCCAGGCTGCACTAAGCGTTTACACAGACCCCAAGCTCCCTCC 1927  
Db 2022 AACCTGATGTTGCAAGATCCAGGCTGCAAAAGCGTACACAGATCTAGTCTCTCC 2081  
QY 1928 GGAACATGTAAGACAGTGTGATGGCCAGAGGCTCTATCTCACCAGAGCAGCAGGG 1987  
Db 2082 GGAACATGTAAGACAGTGTGATGGCCAGAGGCTCTATCTCACCAGAGCAGCAGGG 2141  
QY 1988 ACATCCATCTCGCGCCGACCCGAGAGATTCGGGAGCAGCTTACAGTCCAGTCCG 2047  
Db 2142 ACATCCAGCCAGGCTCCAGCCAGAGAGCCAGGAGCCAGGCACTCTCAGACAGCATCG 2201  
QY 2048 CTGGCCGACCACTCAGGAGGCTTGTGTGAGCAGCAGCTCAGCAACACTACTCTCA 2107  
Db 2202 CAGCGCAGAGACTCAGGCTGGGAGCAGAGCACTCTANTACACTACTCTCA 2261  
QY 2108 AGCGGAAAGATGCTCCAGCTGAAACCGCTCAAGCAGAGAGCAATGATCTCAGC 2167  
Db 2262 GSCGTGAAGATGCTCCAGTGAAGCAGTCAAGTCAGAGAAACCAATGATCTCAGC 2321  
QY 2168 CAGGCGCTGGTGTGAGTCTTATCAGCAGCCACAGCTCCCATCAGCACTATTCCA 2227  
Db 2322 CAGGCGCTGGTGTGAGTCTTATCAGCAGCCACAGCTCCCATCAGCACTATTCCA 2381  
QY 2228 ACAGTGGGCTCAGGCTTCTGACCGATGGAGGTAGTATAGGAGACCTCAGTGCATCG 2287  
Db 2382 ACAATGGATCGAGCTTACTCTGAGTGGCAGTGTAGGAGACCTCAGTGTCTCG 2441  
QY 2288 ATGAAACCCCAATATGAGTCAACCAATTCACCTGCAACACAGCCCTTGTGTCAG 2347  
Db 2442 ATGAAACCCCAATATGAGTCAACCAATTCACCTGCAACACAGCCCTTGTGTCAG 2501  
QY 2348 CCAGGAGAAACCCGCGAGGACCAATGATGGAGCAGTAAACTAGAAAGGCTTAAAC 2407  
Db 2502 CCAGGAGAAATATGAGGAGCAAAATGATGGAGCAAGTAAATAGAAAGGTTGAAC 2561  
QY 2408 AAGTGAATGGAATGTTCCGCGACTGAACCCCATTTCTACCCCTTAAAGCCCTCGGCT 2467  
Db 2562 AAGTGAATGGAATGTTCCGCGACTGAACCCCGTTCCACTTCCAAAGCCCTCGGCT 2621  
QY 2468 CTCTCTCATAGAAATGAGCAGACAGTCCAAACACCTGCAAGCTTGGTGGGCCATGA 2527  
Db 2622 CGCCTCTCATAGAAATGATGAGTCAACAGCAGCTGCAATGTAGGAGGATCCATGA 2681  
QY 2528 CGCTTCTCCGCGAGAGGAGCTCTCTGCGGTGAGGCTCACTGCTGATGATGCTCA 2587  
Db 2682 CTATTCTGCAAAACAGGAACTTTCAAGTACAGATCACTGTTGTGAACGTTGATA 2741  
QY 2588 ACAGAGGAGCAGCAGCAGCAGCAGCTCAGCTCCGCTTACCTGAGCAGCCGCTCCT 2647  
Db 2742 ACAGAGGAGCAGCAGCAGCAGCAGCTCAGCTCAGCTTACTTGTAGCAGCCGCTCT 2801  
QY 2648 CAGGATCTCGCCCTGCTTCTCCAGCCGCTCTCCAGGAGCGCTCAGAGCCGAGGCC 2707

Db 2802 CCGGAATCTCACTTCTCTCCAGCCGGAGGTCAGTGTATGATGCTCTCAGGCTGAGGGAA 2861  
QY 2708 GGCAGCAGAACGTCAGCTGGCGACTCTACAGACCCATCTCCACGACGCTCGCGCC 2767  
Db 2862 GACCACAAATGTGAGGTTGGGACTCTATGACCCCAATTCACAGATGTTCTCGGC 2921  
QY 2768 GTCTCAGCAGAACGACCGAGAGCGGCTGCCACGCTGCTCAGCCTTCAGCCGCGCC 2827  
Db 2922 GCTCCAGTGAAGCAAGCAATGTATGATGACTGCGAGTCTTCTCAGCCTCACCCAGCC 2981  
QY 2828 AGCAGTACCGCTCAGGCGCAAGTACGCGGCTGCCACAGAGGCGCGCGAGCGCC 2887  
Db 2982 AGCAATACAGGCTGAAGTAATATGACAGCTACTGCTGGACCCCAACCTCAC 3041  
QY 2888 TGCCCAACATGAGAGGATGAGCTGAAGACGCGCTGCGTGTCTCGGGGATGCCCTG 2947  
Db 3042 TGCTTAATGAGAGATGAGCCTTAAACAAGATGCACTCTTGGGTGACTGAGGG 3101  
QY 2948 AGCTGGGT---GGCCCTGCTCCAGTTTCAATGCCCCGAGGAGTGCAGCGACGGGAG 3004  
Db 3102 AGTCTGGAGTATCTCCACTGCTCCAGTGAATGCCCTCGAAGAGTAGTGTGGGG 3161  
QY 3005 CCCACGCTACGGCGCGCCACTTGCACCGCAGATGCGCTGGGCCACGCGTGAGGA 3064  
Db 3162 CAATGGTTACAGCAGGAGGATCTTCTATCTCATGATGCTCTAGGAATGGGACGAGA 3221  
QY 3065 GGCCAGCAGCAGCGGTGCGAGAGGCTCGAGGCGCTGCGCTGCTGTCGCGCGCT 3124  
Db 3222 GAGCCAGGATCCGGTAAGATGGTCTCGATAGCCAGTCTTCTAGAGTCCAACTT 3281  
QY 3125 TCAGCAGCTCAGCAGCTGCAACCCCGCGGATGCGCAGTCCGCGGAGAGAGCAGTC 3184  
Db 3282 TCAACAGTCTTAATAGCTTAACTCTCTCTGCTTGGCTCCATCCATGGAACCAACC 3341  
QY 3185 TCGTCTTCAGAAATACGCGCGCGAGGCGGCGAGTCCGCAACTTCCACTGCTCC 3244  
Db 3342 GTGTTCTTCAGAACTGATCCGCTTCGAGGGTGGCTTTTCGCTGGCT---CAGTCCC 3398  
QY 3245 CTGTCTCTCCAGCATCACGAGAGCTCACCTGTGAGTCTCCAGTGGAGCTGATG 3304  
Db 3399 CTGTCTCTCCAGCATCACGAGAGTGTCACTGTGAGTGGAGTGGAGTGGAGTGG 3458  
QY 3305 CCAACTGAACATGAGATGAGTTCCTGCGGAGAGCTGCTGAGTGAATTAATCCAG 3364  
Db 3459 GCAGTTGAATGATGAGTATCTCTGCGAGATGATGCTGAGTACCTGATTCGAG 3518  
QY 3365 ACCAAGCGGTACGAGCAGCTTCCCGAGCGCTCCCGAGCAGCAGCAAGTGC 3424  
Db 3519 ACCAGGCGATGAG---ACCACTTGTGAACAATGTCTAGACAGCAACAATGCAAC 3575  
QY 3425 ACGGCGCGGTGACTTTGACGCGCGCGGCTGCCAGACAGCCAGCTGCGCAGCTTCC 3484  
Db 3576 A-----CAGTGTGGTTTGAAGAAACAATCTCAGAGCTTTGACCGCCCTCCAG 3629  
QY 3485 ATGCCCTCAGCAGCGCTCGCGGAGGAGCAACAGCAGCTGCGCATTCAGTGGAG 3544  
Db 3630 CAAGCAGTCAAGCTGCGAGCTCAGAGGAGCAAGAGAGCTTACCCATTCAGTGGAA 3689  
QY 3545 AAGTCAAGTCCGGAAGCGCGAGCTGCTCTCTCCCAAGTCAAGTGTGGCGCGCGCC 3604  
Db 3690 AGTAACTCAGAAAGTCTGAAATATCTCTTACAAACTGAAATGTAGTCAAGCTCAG 3749  
QY 3605 CTGTGCGCAGACTCGCGCTTTGGGTTCTGCAAGGATGCTGCTCCACCGCAGAAACC 3664  
Db 3750 CAGGCGAGCAAGCTAGAGCTTTGAATGTAATAACAATGATGTTTCAAGCAGCAACT 3809  
QY 3665 CTTGAGGAGCGGCTGCTGGGGGTATCAGACCGCTCGGGGAGAACAGCAACCGCTAG 3724  
Db 3810 TGCAGAGAGC-----AATGTCCCGCAAAATGGGTATCAGAGCTGATGGAACA 3863  
QY 3725 GTGGCCAGCAGCTGATGCTCCACAACAGCCCGGAGTGGCAGAGTGGAAAGCGCT 3784  
Db 3864\*GTGGTCTTATTTTGCAGCAGACACAGGCTTGGCAACGAGGAGCT---AGTCTT 3920

QY 3785 TCCATGAACGCCCTGTAAAGCCCGCAGTATGGAACTGTCTCAACAGGAGCAGCTGG 3844  
Db 3921 TCAGATGACGCCCAATAA-----GCTTACAGTGAAGCATCAGCAGCAACGATGA 3974  
QY 3845 CCCCTGTGCACTCGACGCTGCTGTGTGCGGGATTCAGCCCTCAAGCTGAAGAGCA 3904  
Db 3975 TTTCTGGGCAATGSAATATCTCTGTGCGGTGACAGTTCAAGGGCAGAGTAAGAGCA 4034  
QY 3905 CCCCATGCAAGGAGGGGGCCAGCTGAATTTTCGGCTCCCGGTAGCCCAATGAGT 3964  
Db 4035 GCAACATGCTCGGAGTGGGAACAGCAAAATCTGTGCTATCCCATCAGCTCCAGTGGT 4094  
QY 3965 CAGCTGGCAGCATGGTGAATGGCATGAGAACAGGAGCCAGTGGGACAGGGGTACCTGG 4024  
Db 4095 AAGCCACTGATATGGCAGATGGTATGAGAACAGACATGATGGGACAGGAATATTGG 4154  
QY 4025 CTCACAGCTCCTCGGGACAGCATGAGCACCCGGGGAGGCGCCCGGTACAGAGA 4084  
Db 4155 AAAACCAACTGATTTGGAGATGGCATTCATTAACAGGAGTGAATCAGTCTGGTCCAATGA 4214  
QY 4085 TGCTTGGCAGATTTAGTCTACCTCACACATCAACATCTACCAAGGGCCAGAGCTGCC 4144  
Db 4215 TGCTAGGGCAGGTTAGTCTTACCTCAAAAGCAGCTCTATCAAGGGCCGAGAGTTGTC 4274  
QY 4145 TCCAGGGGCTCACGGCATGGGACAGCGCTCAAGCTTGGCAGTTGTTCAGGGGCTACC 4204  
Db 4275 CACCAGTATTTATACCATTTGTTAAACAGTCTTCAGGTTTGTCAAGTGGCAAGATTACC 4334  
QY 4205 AGCATGTGCGAGCTTTGGGGGAGCAGGCGCCAGGCTATGCCAGGAGCAGACCTTGCTC 4264  
Db 4335 AGCCATGTGCTTAATACAGTGGCAACAGCGGCAAAATGTGTTGAGGAACAACCTGCAC 4394  
QY 4265 TGAGTCAGGACAGCTCAGTCACACAGTCAAGCTCAGGCTCAAGTGAATGATCAAGATGG 4324  
Db 4395 AACAGCAAGGACATGTAGTGTATTAACACAGAGCTGAGGGTAAACACCATTAAGATGG 4454  
QY 4325 AGATGAAGGAGCAGCCCATCCGCTGCTCTAATCTGCAGAAATTAATCTGGTCAGTTCT 4384  
Db 4455 AGATGAAGGTCATCAACAGCTTCTCTCTAATATGAGAATTAATCTGGTCAGTTAT 4514  
QY 4385 ATGACAAACCGTGGGCTTCAAGTCAAGCAAGCAAGCAAGCTGG---TTCATCTCTATTT 4441  
Db 4515 ATGACAAACCGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 4574  
QY 4442 CAGAGCGCAGCTGCTGCTACAGGGA---CCAGCGCAAACTCTGAGTACTTCTTCCC 4498  
Db 4575 CAGAGCTAAGTCTGCTGCGAGGAGTGTCTGCAAACTCTCTGAGCTTCTTCCC 4634  
QY 4499 CAGTGTCTAATCAGGTGACAGCAGTGGAGCTGAGCAGCTGAGCAGCTGGAAGGG 4558  
Db 4635 CAGGGCTAACCAAGTATCAAGCAGCTTGCAGCTTGCAGCAATAGCTAGAGGGTG 4694  
QY 4559 TACAGATGAGTTCGATGCTATGAGCAGTGGGAGCAGCTCAGCCTGATGTGGGGG 4618  
Db 4695 TGCAGATGATTTTGTGCTATCATGATGAGGAGCAGTGTGAGCTTAATTTTCAAG 4754  
QY 4619 CCCTGAGCCCAAGTATCATTCAGAACCTTCCCATAGCTCTCCCGCTCAGCAGCTC 4678  
Db 4755 CCCTGAGCCCAAGTATCATTCAGAACCTTCTCCCGCAATCTCTCAGCTCAGCACTCCC 4814  
QY 4679 GGGGCTCCCTCCCATTCAGTCTGCTCTGATGAGCAGCAGCAGCTGAGCTGAGGGGAC 4738  
Db 4815 GAGCTCTCTTACATTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4873  
QY 4739 ATGAGTCTTGTGCTGAGCTCCCTAGCGGAGAGAAAGAAATTCCTTGCAGTATGCAATAG 4798  
Db 4874 ATGAGTCTTGTGCTGAGCTCCCTAGCGGAGAGAAAGAAATTCCTTGCAGTATGCAATAG 4933  
QY 4799 GCTTTAGGAAAAAAGACTGCAACAA 4825  
Db 4934 ACATGGGAAAAAAGAAAAA 4960



```

RESULT 5
AF231111 5025 bp mRNA linear VRT 18-JUN-2000
LOCUS Coturnix coturnix potential GLI3 (Gli3) gene N- and C-terminal
DEFINITION fragments, complete sequences.
ACCESSION AF231111
VERSION AF231111.1 GI:8571368
KEYWORDS Coturnix coturnix.
SOURCE Coturnix coturnix.
ORGANISM Coturnix coturnix
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Coturnix.
REFERENCE
AUTHORS Borycki, A., Brown, A.M. and Emerson, C.P. Jr.
TITLE Shh and Wnt signaling pathways converge to control Gli gene
activation in avian somites
JOURNAL Development 127 (10), 2075-2087 (2000)
MEDLINE 20233829
PUBMED 10769232
REFERENCE
AUTHORS Borycki, A., Brown, A.M. and Emerson, C.P.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Cell and Developmental Biology, University
of Pennsylvania, 421 Curie Blvd, Philadelphia, PA 19104-6058, USA
FEATURES
source
1..5025
/organism="Coturnix coturnix"
/db_xref="taxon:9091"
/dev_stage="stage HH12 embryo"
gene
1..5025
/ene="Gli3"
misc_feature
200..571
/ene="Gli3"
/ene="Gli3"
/ene="potential coding region for Gli3 N-terminal
fragment; it is not known whether this sequence is
actually translated"
misc_feature
866..4996
/ene="Gli3"
/ene="potential coding region for Gli3 C-terminal
fragment; it is not known whether this sequence is
actually translated; putative protein product contains the
DNA binding domain"
BASE COUNT 1422 a 1346 c 1194 g 1063 t
ORIGIN
Query Match 53.3%; Score 2695.6; DB 5; Length 5025;
Best Local Similarity 74.2%; Pred. No. 0;
Matches 3627; Conservative 0; Mismatches 1154; Indels 109; Gaps 13;
QY 17 TTTTGTGCGAAGAGAGCTGAAGTAATGAGAGACATATGAGAGCCAGTCCACAGCT 76
Db 162 TATTAGTCAAGAGAGCTGAAGTAATGAGAGTCCACATGAGAGGCGAGTCCCATAGCT 221
QY 77 CCAGGACCACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 136
Db 222 CTACACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 281
QY 137 ATGTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
Db 282 ATGTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 341
QY 197 AGACTTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 256
Db 342 AGACCTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 401
QY 257 TCAGCAAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
Db 402 TCAGCAAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
QY 317 AAGAGATCCATGGTCCCTGCCACAGTGGCGGAGCCCTCTGTGCGGTACCGGAGCGG 376
|||||

```





Db 3750 ATCAAGTAAGCTCAGGAAGTTCTGAATATCTCTTACAAATGAATAGTACGAGCT 3809  
QY 3602 CCCTGTGCGGAGACATCGGCGCTTTGGGTTCTGCAACGCGATGTCGTCACCGCAGCA 3661  
Db 3810 CAGCAGGCGAGCAAGCTAGAGCTTTGAACTTGTATACCAACATGATGGTTTCAGCAGCA 3869  
QY 3662 ACCCTTGTAGGAGCGGCTGTGGGGCTATCAGACCCCTCGGGGAGACACACACCCCT 3721  
Db 3870 ACTTGAGAGAAGC-----AATGTCCCCAGCAAAATGGGTATCAGAGCCCTGATGGAAA 3923  
QY 3722 ACCTGTGCGCAGACATCTGATGCTCCACACAGCCCGGAGTGGCAGCAGTGAAGAG 3781  
Db 3924 ACAGTGTCTTATATTTGACAGCAACACAGGCTTGGCAACGAGCAGT---AGTT 3980  
QY 3782 CTTTCATGAACAGCCCTGTGAAGCCCGCGAGTATGGGAACTGTCTCAACAGCAGCAG 3841  
Db 3981 CTTTCAGCATGACAGCCCAATAA-----GCCTTACAGTGAAGAGCATCAGCAGCAACCGA 4034  
QY 3842 TGGCCCTGTGTGACACGAGGCTGTGGTGGCGGATTCAGGCTCAAGGCTCAAGCA 3901  
Db 4035 TGATTTCTGGGCAATGGACAATCTCTGTGGCGTGACAGTTCAAGGGCAGAGCTAAGAA 4094  
QY 3902 GCACCCCATGCAAGGCGGCGGCGAGCTGAATTTGCGGCTGCGGAGTGGCCCAATG 3961  
Db 4095 GCAGCAACATGCTGCGAGTGGGAACACGCAAAATCTGTATCCATCAGTCCAGTG 4154  
QY 3962 AGTCAGTGGCAGCATGTTGAATGGCATGCAACAGCAGCAGCAGTGGCAGCAGGTCAC 4021  
Db 4155 GTCAAGCCACTGATATGGCAGATGGTATGCAAGCAAGCAATGATGGCAGGAAAT 4214  
QY 4022 TGGCTACACAGCTCTCGGCGCAGCATGCAAGCAGCCCGGGGCGAGCCCGGTCAGC 4081  
Db 4215 TGGAAACCAACATGATGGAGATGGCATATTCACAGGAGTGAATAGTCTGGTCCAA 4274  
QY 4082 AGATGCTTGGCAGATGATGCTACCTCACATCAACATCTACCAAGGCGCAGAGCT 4141  
Db 4275 TGATGCTAGGCGAGTGTAGTCTACCTCACAGAGCCTCTATCAAGGCGCAGAGTT 4334  
QY 4142 GCCTGCCAGGGCTCAGGCGAGGCGAGCAGCGCTCAAGCTTGGCAGTGTGAGGGCT 4201  
Db 4335 GTCCACAGTATTTATACATACATGTTGAACAGCTTTCAGTGTGTCAGTGGCAGAGTT 4394  
QY 4202 ACCAGCATGTCCAGCTTTGGGGGCGAGGCGCAGGCTATGCCGAGGAGCAGCTTG 4261  
Db 4395 ACCAGCATGTGTAATTAAGTGGCAACAGAGCGGCAAAATGTTGAGGAACAACTGA 4454  
QY 4262 CTCTGAGTCAGGACAGCTCAGTGACACAAAGTCAGACCTGCGAGGTGAATGTTCAAGA 4321  
Db 4455 CACAACAGAGGACATGTAAGTATTACACAGAGCTGCGAGGTAAACACCATTAAGA 4514  
QY 4322 TGGAGATGAAGGCGACCCCATCCGCTGTGCTTAATCTGCAAGATTAATCTGTGTCAGT 4381  
Db 4515 TGGAGATGAAGGTCAATCAACAGCAGTCTGCTTAATATGCAAGATTAATCTGTGTCAGT 4574  
QY 4382 TCTATGACCAAAACGTTGGGCTTTCAGTCAGCAACAGCAAGAGTGG---TTCAATCTCTA 4438  
Db 4575 TATATGACCAACCATGAGGCTTTAGCCCAACAGCTATGAACAGAGTCTTCTTCTTTG 4634  
QY 4439 TTTTCAGAGCGAGCTGCTCTACAGGGA---CCAGCGGCAAAACTCTGAGTTACTTT 4495  
Db 4635 TTTTCAGAGCTAATGCTGCTCTACAGAGACTGCTACTGCAAACTCTCTGAGACTCTTT 4694  
QY 4496 CCCAGGTGCTAATCAGGTGACAGCAGTGGGACAGCTTCGACAGCCTGACCTGAGAG 4555  
Db 4695 CCCAGGGCTAACCAAGTATCAGCAGGTTGACAGCTTGACAACTAGCCTAGAG 4754  
QY 4556 GGTACAGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4615  
Db 4755 GTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4814  
QY 4616 GGGCCCTGAGCCCAAGTATCATTACAGACCTTTCCATAGCTCTCCCGCCCTCAGCAGC 4675

Db 4815 GAGCCCTCAGCCCAAGCATCATTCAGAAATCTCTCCCGCAATTCCTCAGCGCTCAACA 4874  
QY 4676 CTCGGCGCTCCTCCCATTCCTCAGTCTGTCATGACACACACACACATGGCTATCGGG 4735  
Db 4875 CCGAGCGCTCTTACATTCCTCAG-CTGTGCTGTGAAGCACAACACACATGGCTATGGG 4933  
QY 4736 GACATGAGTCTTCTGCTGACCTCCTAGCGGAAGAAAGCAAAATCTCTTGCAGTTATGCAA 4795  
Db 4934 GACATGAGTCTTCTGCTGACCTCCTAGCGGAAGAAAGCAAAATCTCTTGCAGTTATGCAA 4993  
QY 4796 TAGGCTTTAGGAAAAAAGACTGCAACCAA 4825  
Db 4994 TAGACATGGAAAAAAGAAAAAAGAAAAA 5023

RESULT 6  
AF222990  
LOCUS AF222990 4632 bp mRNA linear VRT 03-MAR-2000  
DEFINITION Gallus gallus GLI3 mRNA, partial cds.  
ACCESSION AF222990  
VERSION AF222990.1 GI:7141287  
KEYWORDS  
SOURCE Gallus gallus.  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 4632)  
Wang, B., Fallon, J.F., and Beachy, P.A.  
Hedgehog-regulated processing of Gli3 produces an  
anterior/posterior repressor gradient in the developing vertebrate  
limb  
Cell 100 (4), 423-434 (2000)  
JOURNAL  
MEDLINE 20155426  
PUBMED 10693759  
REFERENCE 2 (bases 1 to 4632)  
Wang, B. and Beachy, P.A.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (11-JAN-2000) Molecular Biology and Genetics, The Johns  
Hopkins University, School of Medicine, PCTB714, 725 N. Wolfe St.,  
Baltimore, MD 21205, USA

FEATURES  
Location/Qualifiers  
1..4632  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
1..>4632  
/codon\_start=1  
/product="GLI3"  
/protein\_id="AA37273.1"  
/db\_xref="GI:7141288"

CDS  
/translation="MEAGSHSSSTTEKKKVVNSIVKCSNRTDVEKAVASSTTSNDE  
SPGQTYHRRNNAITMQPGGSLKISEPSTSEERASLIKEIHGSIHLPSPV  
PYRGLTMDPRNGMDPHYHPHLOAFHPVPIDARHHEGRTYHYPSIPPLHVS  
ALSSPTVSDLPFTRIISPHRPAASPTSTPHYINPYMDYIRSLHSPSLMSIA  
ARGISPTDPAHAGYSPAYVHOMALLAGORSPYADIIPSAATAGAGALHWEYHAMD  
TRFSPRLSARPSKRRLISPLSDHFDLQTMRTSPNSLVITLNNCSNWSAGSY  
GHLASAIAPALSFTYPTPVSLQMHQIISRQITLGSAGFSPPIHHPAFTPTQR  
PIPGISVLPNVPVSGSPSTOHKPTSESASTGDMHNRKSLKPEDLPSPGA  
GSVOQEGMTPVKEEGDKDEKOEPEVYETNCHWEGCSREFDTQQLVHINNDH  
HGEKKFVRLDCEKQPFKAOYMLVHMRRHTGEKPHKCTPEGCTKAYSRLENL  
TLHSHTGEKPYVCEHEGCKAFNSADRAKHONRTHSNKPYVCKTPGCKRYTDS  
SLRKHVTGHPPEAHVTKQKGDIDPRPPPPRPGSHSQIRSPQQTQGTATGQKDLN  
STTSRECEQKAVKSEKPMTSQSPGGSTCSSESPISNYSNNIGIELTLTGGSV  
GDLVIDETPTMDSTISTATGALQARNNTGKWMQVKLERLKVQNGMLPLNVP  
PPSKAPTLPLPLIGNTSNCSVSGSWTILPNRNLSDITVLNMLNRDSNTSI  
SSAYLSRRSSGISPCFSSRRSPDASOAGPQNVSDVSDPSTDSASRRSSASOC  
DDLPLSLTTPAQOYRLKAKYAATGPPPTPLNMRMSLKTMLLGDREGISIP  
LPYNAPRCSGGANGTSHRLLSHDALNGTNRASDPVMSVNNLUVPRVQHNLSL  
NSFNAPLPPSMERKRLVQNCTHSEGGVFRGSPCPSPISNVTLEAVTMEAGSL  
NDEDLDDVVQYVQVLSNQGTCDHLLNVLDSNMKHHSVVLGNNSPSPDRAPASSQ  
PAGESVSKDLPIOWNEVSGSSDLSPTKLKCSORSAVOQARAGFLYNNMVQOONQ  
RCNVYQNGYQNLNENGSYLOQNTVLGSCASSFGNQPNKPYGESVSRQPMIFGAM  
DSSCGITVQGQKLRSNMPSVSGNQNFHPIASSDQATSWANGQNRNMEQETLQNE

LVGDGTHYGVNOSSPWTIGVSPTSQSSLHQGPQSCPVSHTIGNQSSGLSVAKSYQ  
 PCANYSGNRNRRNLAQQGVSDGNOTYRVNTIKLEMOGQOQFCSNMQNYSGQ  
 LYDQTTGFSHQAMKIGSFEVSPANCLLOETATANSSELLSPCANQYSTVVDLSNS  
 LEGVQIDFAIDIDGSHVLISGALSPSIONLSRNSLTTPRASLTTP"  
 1327 a 1243 c 1058 g 1004 t  
 BASE COUNT  
 ORIGIN

Query Match 53.1%; Score 2685.8; DB 5; Length 4632;  
Best Local Similarity 75.2%; Pred. No. 0;  
Matches 3505; Conservative 0; Mismatches 1112; Indels 42; Gaps 11;

QY	55	ATGGAGCCCGAGTCCACAGAGTCCAGAGCCACTGAAAGAAAAGTTCAGAAATTCCTATA	111
DB	1	ATGGAGCCCGAGTCCACAGAGTCCAGAGCCACTGAAAGAAAAGTTCAGAAATTCCTATA	60
QY	115	GTGAAGTGCCTCCACTCGAAACAGAGATGTGAGCGAGAAAGCGGTTGCCTCCAGCAGCACTTCT	174
DB	61	GTGAATGCTCCCAATCGAAGAGTGTGAGAGAGCTGTTCCTCCAGCAGCACTTCC	120
QY	175	AATGAGGATGAAAGTCTTGACAGAGCTTATCACAGAGAGAGAGAAACGAATCACTATTG	234
DB	121	AATGAGGATGAAAGTCTTGACAGAGCTTATCACAGAGAGAGAGAAACGAATCACTATTG	180
QY	235	CAGCCACAGAATCTCCAGGGCTCGCAAGTCAGTGAAGTCAAGTGAAGACCTTCACATCGAGTGAC	294
DB	181	CAGCCACAGGTGGCAGAGCTTCAGCAGATTCAGTGAAGAGCTTCAACATCTACGAG	240
QY	295	GAGAGGGCCCTATTGATCAAGAAAGAGATCCATGGTGCTCCGCCACAGTGGCGGAGGCC	354
DB	241	GAAGGGCCCTATTAAATAAGAAAGAGATCCATGGATCTATATCGACCTTCCCTGAACCT	300
QY	355	TCGTGCGCGTACCGCGGGACGGTGTTCGATGAGGACCCAGGAATGGTTACATGAGGCC	414
DB	301	TCGTGACCTTATCGTGGGCAATGGTTACTATGAGACCCAGAAATGGTTACATGATGCT	360
QY	415	CACCTACCAACCCTCATCTTTCCCTGCTTCCATCTCCATCTCTGTAACCAATTCATGCCAGA	474
DB	361	CATTACCATCCACTCACTCTTCAAGGGTTCACCCCTCTGTACCAATGATGGCGGA	420
QY	475	CATCATGAGGGCGGTTACCATACGATTCATCTCCGATTCCTCCATTCGATATGACTTCC	534
DB	421	CATCATGAGGGAGCGCTACCAATTATGAACCATCTCCCAITTCCTCTTGCATGTGCTTCT	480
QY	535	GCCTTATCTAGTAGACCCCTAGCATATCGGAGCTGCCTTCATTAGGATCTCCCCACACCGG	594
DB	481	GCCTTATCTAGTAGGCCACATATTCAGATCTTCATTCATTTAGAAATTTCCCCACACAGA	540
QY	595	AACCCCGCTGCTGCTTCGAGTCTCCCTTCAGCGCTCCACATCCCTACATTAATCCCTAC	654
DB	541	AATCCTGCTGCAGCATCAGATCTCCCTTCAGCACTCCTCACCGTACATTAACCCATAC	600
QY	655	ATGGACTATATCGGCTCTTCGACAGAGCCCATCGCTCTCATGATCTCAGCAACCGGT	714
DB	601	ATGGACTATATCGGCTCTTCGACAGAGCCCATCGCTTCCATGATCTCAGCAGCGCGT	660
QY	715	GGCTGAGCCCTACAGATGCGCCCATGAGAGTCCAGCCAGCAGATACATCATCAG	774
DB	661	GGACTCAGCCCAACAGAGCTCCCATGCTGGGCTCAGTCCAGCTGAATATTATCATCAG	720
QY	775	ATGGCCCTGCTAACTGGCCAGCGACCCCTATGACAGACATTAATTCCTTCAGCTGCCACC	834
DB	721	ATGGCTCTGTTGGCAGGCCAGCGACCCCTTAGCAGAGACATCATTCCTTCAGCTGCCACT	780
QY	835	GCCGGCAGCGGGGCAATCCACATGGAAATATCTTCATGCTATGGATAGCAGCAATTCCTC	894
DB	781	GCTGGAGCTGGTCTCTCATATGGAATATCTTCAGGCCATGGATAGCAGCAATTCCTC	840
QY	895	AGCCCCAGGCTGTACGACGAGCGGCGGCGGAAAGCTACACTGTCCATATCACCACTCTCC	954
DB	841	AGTCCGAAGATTATCAGCTAGACCAAGCCGAAAGCTGACGTTGCTATATCCCCCTCATCT	900
QY	955	GATCATAGCTTTGACCTTCAGACCATGATAAGGAGCTTCCCAACTCTCTGGTACGATT	1014

Db	901	GATC	CACAGCTTTGACCTTCAGACCAATGATACGGACATCTCCAAATTCCTTGGTCACAATT	960
Qy	1015	CTCAAT	AAATCCCGCTAGCAGCTCTTCAGCAAGTGGCTCCTATGGTCACCTTATCTGCAAGT	1074
Db	961	CTCAAT	AAATTTCTGTAGCAGTTGGTCAGCCAGTGGTCTTATGGCCACTTATCTGCAAGC	1020
Qy	1075	GCAAT	CAGCCCTGCCCTTGAGCTTTCACCTACTCTTCGCGCCCGCTCTCTCTC- --CACATG	1131
Db	1021	GCAAT	AAGCCCTGCTGTAGTTTCACATACCCCTCTCACACGGTATCCCTCCAGCAAAATG	1080
Qy	1132	CATC	CAGCAGATCCTAAGCCGACACACAGAGCTTAGGTTCAGGCTTTGACACAGCCCTCCA	1191
Db	1081	CATC	AGCAAAATTAAGTTCGTGAGCAAAACCTTGGGTTCTGCTTTGGACACAGCCCTCCT	1140
Qy	1192	CTCAT	CCACCTCCGCCCAACTTTTCAACACAGAGCCCTATTTCAGGGATCCCTACCGTT	1251
Db	1141	CTCAT	CCATCCTGCTCCAACCTTTTCTACCCAGACACTATCCCTGGCATCCCACTGTCT	1200
Qy	1252	CTGA	ACCCGCTCCAGCTCAGCTCCGCGCCTTCTGAGTCCCTCACAG- --AACAAGCCCAAG	1308
Db	1201	CTGA	ACCCCTGTCAGCTCGAGCTTGAGCTTGAGTCCACACAGCACAAACAAACCCACA	1260
Qy	1309	AGTG	AGTCTCGAGTGAAGCAGCACTGGTGACCCGATGCACACACAGAGGTTCCAAGATCAA	1368
Db	1261	AGTG	AGTCTCGAGTGAAGCAGTATCTGGGGATCCAAATGCACAACAAGCCGCTCCAAGATAAAG	1320
Qy	1369	CCCG	ATGAAGACCTCCCAAGCCCAAGGGCTCGGGGACAGGAGCAACAGCCGAAGGAACA	1428
Db	1321	CCGG	ATGAGGACCTCCCAAGCCCGGTGGGGAACGCTGCAGGAACAGCCGAAAGGTTATG	1380
Qy	1429	ACCC	TGTCTCAAGGAGGAGGGACAAAGATGAAGCAACAGGAGCCCTGAAGTCACTCAT	1488
Db	1381	ACCC	CGGTAAAGGAGGAGGACAAAGATGAAGCAACAGGAGCCTGAAGTGGTCTAC	1440
Qy	1489	GAGAC	AAACTGCCACTGGGAAGGCTCGGGAGGAGGTTGACACCCACAGAGCAGCTTGTG	1548
Db	1441	GAGAC	GAACCTGCCAATTTGGGAAGGCTGTTCCCGGGAGTTTGACACACAGGACACAGTAGTG	1500
Qy	1549	CAC	CATATAAATAACGACCATATTCATGAGAGAAGAAGGAGTTCTGTGTCAGTGGCTG	1608
Db	1501	CAT	CATATAAATGACACATACATGGTGAGAAGAAGAAGTTTGTGCGCGATGGCTG	1560
Qy	1609	GACT	GTCTCAAGAGAGAGAAACCTTCAAGCCCAAGTATATGTTGGTAGTGCAATAGAGA	1668
Db	1561	GACT	GTTCGCGGAGCAGAAACCATTCAAAGCCCAAGTATATGCTGGTGGTCCACATGAGG	1620
Qy	1669	AGAC	CACAGGGCGAGAAGCCTCACAAATGCACCTTTGAAGGTTGCACAAAGGCGCTACTCG	1728
Db	1621	AGAC	ATACAGGGGAAAAGCCACACAATATGCACTTTGAAGGTTTGTACAAAGGCGCTACTCC	1680
Qy	1729	AGAC	TAGAAACTTTGAAAACACACTTGAGATCTCACCTGCAGAGAACAACATAGCTCTGT	1788
Db	1681	AGAC	TAGAAACTTTGAAAACACACTTGAGATCTCACCTGCAGAGAACAACATATGTTGTG	1740
Qy	1789	GAGC	AGAAAGGTTTCTCAATGCGCTCTGATCGCGCCCAACACACAAAC	1848
Db	1741	GAAC	ACGAGGCTGCAACAAGCTTTCTCCAATGCATCTGCAGGSCCAGCACCAAAAC	1800
Qy	1849	AGAA	CGCAATTCGAATGAGAAACCATATGTGTGCAAAATCCCAAGCTGCACTAAGCGTTAC	1908
Db	1801	AGGAC	TCATTCAAATGAAAACCCATATGTTTGCAAGATCCCAAGCTGCACAAAGCGCTAT	1860
Qy	1909	ACAG	ACCCAGCTCCTCCGGAAAACATGTGAAGACAGTGCATGGCCACAGGCTCATGTCT	1968
Db	1861	ACAG	ACCCCTAGTTCTCTCGGAAAACATGTGAAGACTGTGTGATGGCCAGAGCACAGCTC	1920
Qy	1969	ACCA	AAGACGCGAGGGGACATCCATCCTCGGCGCCACCCCGAGAGATTTCCGGCAGC	2028
Db	1921	ACCA	AGACGAGAGACATCCACCCAGGCGCTCCACCAACAGACCCAGCCAGGCAGC	1980
Qy	2029	CATT	CACAGTCCAGGTCGCTTGGCGACCGCACTCAGGAGAGCCCTTGGTGACACAGGAC	2088
Db	1981	CAC	TCTCAGACAGATCACAGGCCAGCAGACTCAGGTGGCGACTGGGGGACAGAGGAC	2040

QY 2089 CTCAGCAACACTACCTCAAGCGGAGAGATGCTCCAGGTGAACACCTCAAGCAGAG 2148  
Db 2041 CTCACAGCACACTACCTCAAGCGGTGAAGATGCTCCAGGTGAACACCTCAAGTCAGAG 2100  
QY 2149 AAGCCAAATGACATCTCAGCAAGACCTGGTGTGCTAGTCTTATGAGCAGGCAACAGTCC 2208  
Db 2101 AAACCAATGACATCTCAGCAAGACCTGGTGTGCTAGTCTTATGAGCAGGCAACAGTCC 2160  
QY 2209 CCATCAGCAACTATPCCAAAGTGGGCTCGAGCTTCTCTGACCGATGGAGGTAGTATA 2268  
Db 2161 CCATCAGCAACTATPCCAAAGTGGGCTCGAGCTTCTCTGACCGATGGAGGTAGTATA 2220  
QY 2269 GGAGACTCAGTCCCATGATGAACCCCAATCATGAGTCTCAACCATTTCCACTCAACC 2328  
Db 2221 GGAGACTCAGTCCCATGATGAACCCCAATCATGAGTCTCAACCATTTCCACTCAACC 2280  
QY 2329 ACAGCCCTTGTCTTGAAGCCAGGAGAAACCCGAGGAGCAAAATGGATGAGCAGCTA 2388  
Db 2281 ACAGCACTTGGCTTACAGCCAGGAGGAATATGACAGGAGCAAAATGGATGAGCAGCTA 2340  
QY 2389 AAACCTAGAAGGCTAAACCAAGTGAATGTTTCCCGGACTGAACCCCAATTTCAACC 2448  
Db 2341 AAATAGAAGGCTTAACCAAGTGAATGTTTCCCGGACTGAACCCCAATTTCAACC 2400  
QY 2449 CCTAAAGCCCTCGCGTCTCTCTCATAGGAATGGACACAGTCCCAACACACCTCG 2508  
Db 2401 TCCAAAGCCCTCGCGTCTCTCTCATAGGAATGGACACAGTCCCAACACACCTCG 2460  
QY 2509 AGCTTGGGTGGGCCATGAGCTTCTCCCGGAGGAGGAGGAGTCTCTGGGTGGAGCTC 2568  
Db 2461 AGTGTAGGAGGATCCATGACTATTCGCCAACAGGAATGAATTTCTAGTACAGACATC 2520  
QY 2569 ACTATGCTGAACATGCTCAACAGAGGAGACAGGAGGAGGAGGAGGAGGAGGAGGAG 2628  
Db 2521 ACTGTGTGAACATGCTTAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580  
QY 2629 CTGAGCAGCGCGCTCTCAGGATCTCGCCCTGCTTCTCCAGCGCGCTCCAGCGAG 2688  
Db 2581 TTGAGCAGCGCGAGTCTTCTGGAATCTCAGCTTCTTCCAGCGCGGAGGTTCCAGTAT 2640  
QY 2689 GCCTCAGAGCGCGAGCGCGCGGAGAGCTGAGGAGTGGGAGTCTCTACGACCCCATC 2748  
Db 2641 GCCTCAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700  
QY 2749 TCACAGCGCGCTCGCGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2808  
Db 2701 TCAACGATGCTCTCGCGCTCAAGTGAAGCAAGCAATGTGATGACCTGCAAGTCTT 2760  
QY 2809 CTGAGCTCAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2868  
Db 2761 CTGAGCTCAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820  
QY 2869 GGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2928  
Db 2821 GGACCCCACTCACTGCTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880  
QY 2929 CTGCTGGGAGTCCCTCGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 2985  
Db 2881 CTCTGGGTGACTGCAAGGAGTCTGGAATATCTCCACTGCTCCAGTGAATGCCCTTAGA 2940  
QY 2986 AGGTGAGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3045  
Db 2941 AGATGTAGTGGTGGGCAATGTTTACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3000  
QY 3046 CTGGCCAGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3105  
Db 3001 CTAGGAATGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3060  
QY 3106 CTGCTCGTGGCGCTTACAGAGCTCAGAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 3165  
Db 3061 GTCCCTAGAGTCCAACTTCAACAGTCTTAATAGCTTTAACTCTCTCTGCTTGGCTTCCA 3120

QY 3166 TCCCGGAGAGCGAGTCTGCTTCAAGATTTACAGCGCGCGGAGGCGGCGAGTCC 3225  
Db 3121 TCCATGGAAAGCGAGACTTGTCTTCAAGAACTGTACCAATTTCTGAGGCTGGTGTTC 3180  
QY 3226 CGAAATTTCCACTCTGCTCCCTCTCCAGCATCAACGAGAACTCAACCTCGGAGTCC 3285  
Db 3181 CGTGGCTT---CAGTTCCTCTCTCCAAAGCATCAGCGAGAACTCAACCTCGGAGCT 3237  
QY 3286 CTGACCATGGAGCTGATGCGCAACCTGAACGATGAGATTTCTTCCGCGGAGGAGTGGTG 3345  
Db 3238 GTTACGATGGAGGAGGCTGGCAGTTTGAATGATGAGGATTTCTTCCGAGATGATGTGGTT 3297  
QY 3346 CAGTATTTAAATTTCCAGAACCAAGCAGGATACGAGCAGCACTTCCCGAGCGCTCCCG 3405  
Db 3298 CAGTATCTGAATTTCCAGAACCGGCGAGCTGGG---ACCATTGTTGAACAATGTCTTA 3354  
QY 3406 GACGACAGCAAGTTCGCCAGCGGCGGCTGACTTTGACGCGCGGCGGCTGCGCAGACAGC 3465  
Db 3355 GACAGCAACAAATGCATCA-----CAGTGTGGTGTAGGAAACAACAATCCGAGCAGC 3408  
QY 3466 CACGCTGCCAGCAGTTCCTATGCTCGAGCAGCGCTGCCCGGAGGCGGAGCAACACGAC 3525  
Db 3409 TTTGATCAGCGCTTCCAGCAAGCAGTCAAGCTCGAGCTCAGAGGTAAAGCAAGTGCAC 3468  
QY 3526 CTGCCCATTCAGTGAACGAACTCAGCTCCGAAAGCGCGGAGCTGCTCTCTCAAGCTC 3585  
Db 3459 TTACCATTCAGTGAATGAAGTGAAGTCTCAGAACTTCTGACTTATCTCTTCAAACTG 3528  
QY 3586 AAGTGTGGCGCGCGCTGCTGCCGAGACTCGCGCTTTGGGTTTGTGAAGCGGATG 3645  
Db 3529 AAATGTAGTCAGCGCTCAGCAGTGCAGCAGCTAGACATTTGGAGCTGATATAACAATG 3588  
QY 3646 GTCGTCACCGCGCAACCCCTTGGAGGAGCGGCGCTGCTGGGCGGCTATCAGACCTCGG 3705  
Db 3589 ATGTTCAACGAGCACTTGCAGAGGCTGCAAGAGCT-----AATGTGTACCAAGAAATGGTAT 3642  
QY 3706 GAGAACAGCAACCCCTACGCTGGCGCAGCAGCTTGTGTCTCCACAGCAGCGCGGAGT 3765  
Db 3643 CAGAACCTGTAGGAAACAATGTTCTTATGTTTGCAGCAGACACCGTCTTGGCAGC 3702  
QY 3766 GGCACCTGTGAAGCGCTTCCATGAACAGCCCTGTAAAGCGCGCGAGTGTGAAGCTGT 3825  
Db 3703 GGAGCAGT---AGTCTTTCGGCATGCAAGCCCAATAA-----GCCTTATGTTGAAAGC 3753  
QY 3826 CTCACAGCGGAGCAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3885  
Db 3754 CTCACAGCGCAACCTTATGTTTGGGCAATGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3813  
QY 3886 GCTCAAGCTGAAGCAGCAGCCCATGCAAGGAGCGGCGGCGAGCTGAATTTGCGGCTG 3945  
Db 3814 GGGCAGAGCTGAAGAGCAGCAACATGCTGAGTGGAGACCAAGCAAAATTTGGTCTAT 3873  
QY 3946 CGGTAGCGCAATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4005  
Db 3874 CCATCGCATCCAGTGATCAAGCCCAAGTATGGCAATGGATGCAAAACAGGAATGTG 3933  
QY 4006 GTGGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4065  
Db 3934 ATGGAACAGGAATATTTGCAAAACGAACTGTTGGAGATGGCAATTTATTTACGAGGAGCT 3993  
QY 4066 GGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4125  
Db 3994 AATCAGTCCAGTCCAAATGACACTAGGCGAGGTTAGTCTTCTTCTTCTTCTTCTTCTTCT 4053  
QY 4126 CAAGGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4185  
Db 4054 CAGGCGCAGCAGAGTGTCTTCCCGTATCTCATACCATTTGTTAACTGCTTTCAGTGTG 4113  
QY 4186 GCAGTGTGAGGGCTTACAGCCATGTCAGCTTTGGGCGCAGCAGCGGCGGAGCTATG 4245  
Db 4114 TCAGTGGCGCAAGATTTACAGCCATGTGCTAATTTACAGTGGCAACAGCAGCAAAATGTG 4173  
QY 4246 CCGAGGAGCGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4305

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Jun 4, 2002 this sequence version replaced gi:3419842.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [saplens@watson.wustl.edu](mailto:saplens@watson.wustl.edu)  
----- Summary Statistics  
Center project name: H\_GS539F22  
-----

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

**SOURCE INFORMATION:**  
This clone is from the first BAC library from Genome Systems, Inc.  
(<http://www.genomesystems.com>).

**NEIGHBORING SEQUENCE INFORMATION:**  
The clone sequenced to the left is GSI-512I21, 200 bp overlap; the clone sequenced to the right is GSI-489L14, 200 bp overlap. Actual start of this clone is at base position 1 of GSI-539F22; actual end is at base position 189279 of GSI-539F22.

This clone contains polymorphisms with GS512I21.

FEATURES	source
mus_c10ne_contains_polyomorphisms	Location/Qualifiers
	1. .189279
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7p12-p14"
	/clone="GSI-539F22"
	/clone_lib="GSBAC1"
	1. .98
repeat_region	/rpt_family="MER1_type"
repeat_region	694. .920
repeat_region	/rpt_family="L2"
	1475. .1776
repeat_region	/rpt_family="Alu"
	2358. .2647
repeat_region	/rpt_family="Alu"
	2981. .3255
repeat_region	/rpt_family="Alu"
	5513. .5671
repeat_region	/rpt_family="MER1_type"
	6473. .6548
repeat_region	/rpt_family="Achobo"
	7280. .7435
repeat_region	/rpt_family="MER1_type"

```

repeat_region 7615..7974
/rpt_family="ERV1"
repeat_region 8109..8292
/rpt_family="MER2_type"
/rpt_family="MER2_type"
repeat_region 8480..8566
/rpt_family="L1"
/rpt_family="L1"
repeat_region 10627..10933
/rpt_family="Alu"
/rpt_family="Alu"
repeat_region 11661..11760
/rpt_family="L2"
/rpt_family="L2"
repeat_region 12249..12566
/rpt_family="Alu"
/rpt_family="Alu"
repeat_region 12710..13001
/rpt_family="L1"
/rpt_family="L1"
repeat_region 13120..13307
/rpt_family="L1"
/rpt_family="L1"
repeat_region 14383..14645
/rpt_family="L1"
/rpt_family="L1"
repeat_region 15037..15350
/rpt_family="Alu"
/rpt_family="Alu"
repeat_region 15373..15514
/rpt_family="MER1_type"
/rpt_family="MER1_type"
repeat_region 15632..15780
/rpt_family="L1"
/rpt_family="L1"
repeat_region 15851..16028
/rpt_family="Alu"
/rpt_family="Alu"
repeat_region 16038..16325
/rpt_family="Alu"
/rpt_family="Alu"
repeat_region 16346..16596
/rpt_family="L1"
/rpt_family="L1"
repeat_region 16822..16977
/rpt_family="L1"
/rpt_family="L1"
repeat_region 17040..17436
/rpt_family="ERV1"
/rpt_family="ERV1"
repeat_region 17703..17913
/rpt_family="L1"
/rpt_family="L1"
repeat_region 18138..18275
/rpt_family="MIR"
/rpt_family="MIR"
repeat_region 20452..20602
/rpt_family="MER1_type"
/rpt_family="MER1_type"
repeat_region 20749..21021
/rpt_family="Alu"
/rpt_family="Alu"
repeat_region 21056..21221
/rpt_family="MER1_type"
/rpt_family="MER1_type"
repeat_region 23764..23938
/rpt_family="MIR"
/rpt_family="MIR"
repeat_region 24479..24768
/rpt_family="L2"
/rpt_family="L2"
repeat_region 26480..26683
/rpt_family="L1"
/rpt_family="L1"
repeat_region 26695..27279
/rpt_family="ERV1"
/rpt_family="ERV1"
repeat_region 27312..27433
/rpt_family="ERV1"
/rpt_family="ERV1"
repeat_region 27571..27879
/rpt_family="Alu"
/rpt_family="Alu"
repeat_region 28055..28171
/rpt_family="L1"
/rpt_family="L1"
repeat_region 28207..28861
/rpt_family="L1"
/rpt_family="L1"
repeat_region 28862..29624
/rpt_family="L1"
/rpt_family="L1"
repeat_region 30675..31304
/rpt_family="MER2_type"
/rpt_family="MER2_type"
repeat_region 33346..33499
/rpt_family="L2"
/rpt_family="L2"
repeat_region 33691..33741
/rpt_family="L2"
/rpt_family="L2"
repeat_region 34405..34701
/rpt_family="Alu"
/rpt_family="Alu"
misc_feature 35742..36220
/note="similar to EST AA513919 (NID:g2252340) ng67d09.s1"
repeat_region 36910..36999

```

```

repeat_region /rpt_family="MER2_type"
37118..37375
/rpt_family="MER2_type"
37383..37436
/rpt_family="MaLR"
37554..37754
/rpt_family="MIR"
37768..38169
/rpt_family="ERV1"
38569..38756
/rpt_family="ERV1"
38819..39057
/rpt_family="MaLR"
39158..40125
/rpt_family="ERV1"

Query Match 50.6%; Score 2557.6; DB 9; Length 189279;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2580; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 2456 CCCCTGCGGTCTCTCTCATAGGAATGGCACACAGTGCACACACCTGCAGCTTGG 2515
Db 176349 GCGATGGTCTCTCCCTTTCTCTGTCTCCAGGCACACAGTGCACACACCTGCAGCTTGG 176290

QY 2516 GTGGSCCATGACGCTTCTCCGGSCAGAGGACCTCTCTGGGGTGAGCGTCACTATGC 2575
Db 176289 GTGGSCCATGACGCTTCTCCGGSCAGAGGACCTCTCTGGGGTGAGCGTCACTATGC 176230

QY 2576 TCAACATGCTCAACAGAGGACAGCAGCGCCAGCACCATCAGTCTGGCTTACCTGAGCA 2635
Db 176229 TCAACATGCTCAACAGAGGACAGCAGCGCCAGCACCATCAGTCTGGCTTACCTGAGCA 176170

QY 2636 GCCGCGCTCTCAGGGATCTCGCCCTCTCTTCAGCCGCGCTTCCAGCAGGCGTCA 2695
Db 176169 GCCGCGCTCTCAGGGATCTCGCCCTCTCTTCAGCAGGCGCTTCCAGCAGGCGTCA 176110

QY 2696 AGGCGAGGCGCGCGCAGAACGTCGAGCGTGGCGACTCTACAGCCCATCTCCACCG 2755
Db 176109 AGGCGAGGCGCGCGCAGAACGTCGAGCGTGGCGACTCTACAGCCCATCTCCACCG 176050

QY 2756 AGGCGAGGCGCGCGCAGAACGTCGAGCGTGGCGACTCTACAGCCCATCTCCACCG 2815
Db 176049 AGGCGAGGCGCGCGCAGAACGTCGAGCGTGGCGACTCTACAGCCCATCTCCACCG 175990

QY 2816 TCACGCGCGCGCGCAGGATCTCCAGCGTCAAGCGCAAGTACGCGGCTGCCACAGGCGCGC 2875
Db 175989 TCACGCGCGCGCGCAGGATCTCCAGCGTCAAGCGCAAGTACGCGGCTGCCACAGGCGCGC 175930

QY 2876 CGCGAGCGCGCGCGCAGGATCTCCAGCGTCAAGCGCAAGTACGCGGCTGCCACAGGCGCGC 2935
Db 175929 CGCGAGCGCGCGCGCAGGATCTCCAGCGTCAAGCGCAAGTACGCGGCTGCCACAGGCGCGC 175870

QY 2936 GGGATGCGCTCGAGCGTGGCGCTTCCCTTCCAGTTTCATGCCCGAGGAGTGCAGCG 2995
Db 175869 GGGATGCGCTCGAGCGTGGCGCTTCCCTTCCAGTTTCATGCCCGAGGAGTGCAGCG 175810

QY 2996 ACGGGGAGCGCGCGCTACGGGGCGCGCGCTTCAGCGCGCGCGCTGCCAGCGCGCGC 3055
Db 175809 ACGGGGAGCGCGCGCTACGGGGCGCGCGCTTCAGCGCGCGCGCTGCCAGCGCGCGC 175750

QY 3056 CGGTGAGGAGCGCGCGCGCTTCCAGCGCGCGCGCTTCCAGCGCGCGCGCTGCCAGCGCGCGC 3115
Db 175749 CGGTGAGGAGCGCGCGCGCTTCCAGCGCGCGCGCTTCCAGCGCGCGCGCTGCCAGCGCGCGC 175690

QY 3116 TGCCGCGCTTCAGCAGCGCTCAGCAGCTCAACACCGCGCGCGCTGCCAGCGCGCGC 3175
Db 175689 TGCCGCGCTTCAGCAGCGCTCAGCAGCTCAACACCGCGCGCGCTGCCAGCGCGCGC 175630

QY 3176 AGCGAGCTCTGCTTCCAGAAATACAGCGCGCGCGCGCGCGCTGCCAGCGCGCGC 3235
Db 175629 AGCGAGCTCTGCTTCCAGAAATACAGCGCGCGCGCGCGCGCTGCCAGCGCGCGC 175570

QY 3236 ACTCGTCCCGCTCTCTCTCCAGCATCCAGAGAGCTACCGTGGAGTCCCTGACCATGG 3295

```





Copied from PCI US2003092 on 04-03-2004





4613	CGGGGGCCCTGAGCCCAAGTATCAATTCAGAACCTTTCCATAGCTCCTCCCGCCTCACCA	4672
Db		
4812	CTGGAGTTCTTAAGTCCAAAGTATCATCCGAATCTCTCCGTAATTTCTTCACGCTTTACCA	4871
Db		
4673	CGCCTCGGGCGCCCTCCCATTCGCCAGTCGCTGCCATGAGCAGCACCAACATGGCTATC	4732
QY		
4872	CTTCCTCGAGCATCTCTGACATTTCCAG-CCATACCTGTTAGTACATCTAACTAGGCTATT	4930
Db		
4733	GGGACATGAGTTCTTTGCTGACCTCCCTAGCGGAAGAAAGCAAAATTCCTTGCAGTTATG	4792
QY		
4931	GGGGACATGAGTTCTTTGCTGACCTCACTTGCGAAGAAAGCAAGTTTCTTGCTATTATG	4990
Db		
4793	CAATAGGCTTTAGGAAARAAAGACTGCACACCAACGGAATCAATAGGAGTTGAGAGATT	4852
QY		
4991	CAATAACTTTAGAGCGAGAGAANAATGAACAATGAGCAAGCCCGTA-----	5035
Db		
4853	AAACTGACTTTGTTTGGTGTGTTTTTTAGTTCTGTATGTTATTTTAGCAATCTCATCTCA	4912
QY		
5036	-----ACTCATTTGTATGTGTACGTA- TTCAGCAATCTAATCCCA	5076
Db		
4913	CCTAACTGAGATGTTTTCAAATTATATTCCTTTTATGGAAAAGGACTCGAAAAACCCTA	4972
QY		
5077	AACAAATGGATGTATTCGAAGTACATTCCTTTTATGGAATAGGACTT--AAACCCCTA	5133
Db		
4973	AGTATTCTTAGGAGAAAGCTGCTCCATTTTCAGTTTTCGAATCAGTATTGTTACACTCAA	5032
QY		
5134	AGTGTATAGGAGAAATAGTCTCCATTTCAAGTTTAGACAGTATTGTCNACCCAA	5193
Db		
5033	ACCACCTCTTTTTA	5047
QY		
5194	ACAATGTTGCTTTTA	5208
Db		

```

RESULT 9
AC026814/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC026814
1
MUSKINGHAM CHROMOSOME 13 CLONE RP23-377A8 STRAIN C57BL6/J,
WORKING DRAFT SEQUENCE, 8 UNORDERED PIECES.
AC026814
1
GI:7321459
HTG: HTGS_PHASE1; HTGS_DRAFT.
MUS MUSCULUS.
MUS MUSCULUS
EUKARYOTA; METAZOA; CHORDATA; CRANIATA; VERTEBRATA; EUTELEOSTOMI;
MAMMALIA; EUTHERIA; RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
1 (BASES 1 TO 132169)
BECKSTROM-STERNBURG,S.M., BENJAMIN,B., BOUFFARD,G.G.,
DIETRICH,N.L., EAGLE,W.O., GAN,W., GUPTA,J., HO,S.-L., HUANG,M.C.,
IDOL,J., LEE-LIN,S.-Q., MADURO,Q.L., MADURO,V.B., MASTRIAN,S.D.,
MCCLOSKEY,J.C., MORSE,E., OJODU,M.A., PEARSON,R., STANTROP,S.,
SUMMERS,T.J., THOMAS,J.W., THOMAS,P.J., TONGSON,E.E.,
TOUCHMAN,J.W., FRAN,J.F., VOGT,J.L., WALKER,M.A., WETHERBY,K.D. AND
GREEN,E.D.
NISC MOUSE SEQUENCING INITIATIVE
UNPUBLISHED
2 (BASES 1 TO 132169)
GREEN,E.D.
DIRECT SUBMISSION
SUBMITTED (24-MAR-2000) NIH INTRAMURAL SEQUENCING CENTER, 8717
GROVEMONT CIRCLE, GAITHERSBURG, MD 20877, USA
----- GENOME CENTER
CENTER: NIH INTRAMURAL SEQUENCING CENTER
CENTER CODE: NISC
WEB SITE: http://www.nisc.nih.gov
CONTACT: nisc\_mouse@nhgri.nih.gov
----- PROJECT INFORMATION
CENTER PROJECT NAME: XF
CENTER CLONE NAME: 377A08
----- SUMMARY STATISTICS
SEQUENCING VECTOR: PLASMID; N/A; 100% OF READS
CHEMISTRY: DYE-TERMINATOR BIG DYE; 100% OF READS
ASSEMBLY PROGRAM: PHRAP; VERSION 0.990319

```

Consensus quality: 128527 bases at least Q40  
 Consensus quality: 129561 bases at least Q30  
 Consensus quality: 130446 bases at least Q20  
 Insert size: 126000; agarose-fp  
 Insert size: 133000; pulse-field-gel  
 Insert size: 132169; sum-of-contigs  
 Quality coverage: 10.20x in Q20 bases; agarose-fp  
 Quality coverage: 8.40x in Q20 bases; pulse-field-gel  
 Quality coverage: 9.72x in Q20 bases; sum-of-contigs  
 -----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1  
 \* 2075: contig of 2075 bp in length  
 \* gap of unknown length  
 \* 2076 10396: contig of 8321 bp in length  
 \* gap of unknown length  
 \* 10397 14664: contig of 4268 bp in length  
 \* gap of unknown length  
 \* 14665 26928: contig of 12264 bp in length  
 \* gap of unknown length  
 \* 26929 59135: contig of 32207 bp in length  
 \* gap of unknown length  
 \* 59136 81924: contig of 22789 bp in length  
 \* gap of unknown length  
 \* 81925 103946: contig of 22022 bp in length  
 \* gap of unknown length  
 \* 103947 132169: contig of 28223 bp in length.

#### FEATURES

Location/Qualifiers  
 1..132169  
 /organism="Mus musculus"  
 /strain="C57BL6/J"  
 /db\_xref="taxon:10090"  
 /chromosome="13"  
 /clone="RP23-377A8"  
 /clone\_lib="RPC1 mouse BAC library 23"  
 39642 a 26936 c 27202 g 38378 t 11 others

Query Match 32.2%; Score 1625.2; DB 2; Length 132169;  
 Best Local Similarity 78.7%; Pred. No. 6.4e-294;  
 Matches 2025; Conservative 0; Mismatches 518; Indels 29; Gaps 6;

2485 GGCACACAGTCCCAACACACCTGACGTGGGTGGGGCCGATGACGCTTCTCCCGGGCAGA 2544  
 90580 GGCACACAGTCCCAATAAACACTATATAGCTGGGTGGGGCCGCGGACCTCTCCCGAGCAGA 90521  
 2545 AGCGAGCTCTCTGGGTGGAGCTCACTATGCTGAACATGCTCAACAGAAGGACAGCAGC 2604  
 90520 AGTGACCTGTAGGTGTAGACTTCAAGTGTGACAGTGTGACACTCAACAGGAGAGACAGCAAT 90461  
 2605 GCGAGCACCATCAGCTCGGCCCTACCTGAGCAGCCCGCTCTCTCAGGGATCTCGCCCTGC 2664  
 90460 ACCAGCACTATCAGCTCTCGCTACCTGAGCAGCCGACAGTCTCGGGCATCTCCCGCTGC 90401  
 2665 TTCTCCAGCGCGCTCCAGCAGGCGGTCAACAGCGCGAGGGCGCGCCGCGCAGACGTGAGC 2724  
 90400 TTTTCCAGCGCGAGGTCCAGCAGGCGGTCCAGGCTGAAGGCGGACCCCGCAATGTAGT 90341  
 2725 GTGGCGAGCTCTACGACCCCTCTCCACGAGCGCTCGCGCGCTCCAGCAGGACCCAGC 2784  
 90340 GTGGCTGACTCTCAGATCCCATCTCCACAGATGCTTCAAGAGGTCCAGCGAGGCGCAGC 90281  
 2785 CAGAGCGAGCGGCTGCCAGCGCTCTCAGCTTCAGCGCTTCAGCGCGCGCCGCGCAGTACCGCTCAAG 2844  
 90280 CAGGTGAGCGGCTGCCAGCTCTGCTCAGCTCTCAGCTCTCAGCGCTCAGCGCTACCGCTTAAG 90221  
 2845 GCGAAGTACCGGCTGCCACAGAGGGCGCGCGCGCGCTGCCCGCGCGCTGCCCGCGCGAGG 2904

Db 90220 GCCAAGTATGACAGTGCACAGGTGGCCACACCCACACCTTGCCTCCACATGAGAGG 90161  
 QY 2905 ATGAGCCTGAAGACGCGCTGGCCCTGCTCGGGGATGCCCTCGAGCCTGGGTGGCCCTG 2964  
 Db 90160 CTGAGCCTGAAGACCAAGATGGCTTGTAGGGGAAGGAGGAGTCTGGGGTGAACCTG 90101  
 QY 2965 CTTCCAGTTTCATGCCCGAGGAGGTGAGCAGCAGCGGGAGCCACGCTACGGCGCGCGC 3024  
 Db 90100 CCTCAGTCCATCTCTCGAAGATGAGTGTGGGAGGTACACATACAGAGGCGGT 90041  
 QY 3025 CACCTCAGCGCGCAGATGCGCTGGGCCACGCGTGTAGGAGGGCCAGCCCGGTGCGG 3084  
 Db 90040 CACCTGATGCTCATGATGCACTAGCGAAGAGTGTAGGAGAGCCAGCGCTGTGAGG 89981  
 QY 3085 ACAGGCTCCGAGGCGCTGGCCCTGCTGCTGTGCGCGCGCTTCCAGCAGCTTCAGCAGTGC 3144  
 Db 89980 ACAGTCTCGGAGAACATGTCTACCTTGCAGAGGTTCACGCTTCAGCAGCTCAATAGTTT 89921  
 QY 3145 AACCCCGCGGATGGCCACGCTCCGCGGAGAGCGCAGTCTCGTCTTCAGAAATTACAGC 3204  
 Db 89920 AATCTCCAAATTTGCTCCATCGTGTGAAAGAGGTAGCTTAGTCTGCAAAATTATACC 89861  
 QY 3205 CGGCCCGAGGCGCGCAGTCCCGAAACTTCCACTGCTGCTCCCTGCTCTCCAGCATCACC 3264  
 Db 89860 CGGCAGGAGAGCAGCAACCCCGGTACTTCCAGGATCCCTTGCCTCCAGCATCACA 89801  
 QY 3265 GAGAACGTACCTCGTGGAGTCCCTGACCATGAGCGCTCATCCCACTGAAAGATGAGAT 3324  
 Db 89800 GAGAATGTTCCTCGGAGGCTCTGACCATGATGCTGATGCTAACTTGAATGATGAAGAC 89741  
 QY 3325 TTCTGCGCGAGCAGCTGGTGCAGTATTTAAATTCAGAACCAAGCAGGCTGAGGATGAC 3384  
 Db 89740 CTCCTGCCAGATGATGTGTACAGTATTTAAATTCAGAACCAAGCAGGATGAGGAC 89681  
 QY 3385 CACTTCCCAGCGCCCTCCCGGACGACAGCAAGTGCCTCCAGCGCGCGGTGACTTTGAC 3444  
 Db 89680 CAGTTGACAGAGTGGCATCTCTGAAGACAGCAAGTAGCCATGAGCCAGAGGACTTTGAC 89621  
 QY 3445 GCGCCCGGCTGCCAGACAGCCAGCTGGCCAGCAGTTCCTCCAGTCCCTCGAGCAGCCCTGC 3504  
 Db 89620 TTACAGGGCTGCCAGACAGTATGTTGGCCAGGATACCCAGCTCTGGAGAGCCCTGC 89561  
 QY 3505 CCCGAGGCGAGCAAAACCGACCTGCCCATTTCACTGGAACCAAGTCACTCCGGAAGCGCC 3564  
 Db 89560 TCTGAGGCGAGTAAACTGATTTGCCCATCCAGTGAAGACAGGTCAGTTCTGGAACCTCT 89501  
 QY 3565 GACCTGTCTCTCCAGCTCAAGTGTGGCGCGCGCGCGCT---GTGCGCGAGACTCGC 3621  
 Db 89500 GATCTGCTATCTCTCAAGCTGAAGTGTGGTCAAGCAGCGCTAGCGCACAGCAGCTCGG 89441  
 QY 3622 GCCTTTGGTTCGCAACGGCATGCTGCCACCCGAGAACCCCTTGA-----GGAGC 3675  
 Db 89440 GGTTTGGGCTATACAACAATATGTTGTTACCCACATATCTGTGGAAGTTGGCACT 89381  
 QY 3676 GGGCTCTGGGGCTATCAGACCTCGGGGAGAACCAACCCCTACGGTGGCCCGCAGAG 3735  
 Db 89380 GGCCAGCTGGGGATATCAGACCTCGGGGAGAAATAGCAGTACCTACATGCCCCAGAG 89321  
 QY 3736 CACTTGTATGCTCCACAACAGCCCGGGAAGTGGCACCACTGGAACGCCCTTCCATGAACAG 3795  
 Db 89320 CACTTTTCAATCCACAGCTGGGGATGGACTTGGCACCAATGGAATACTTTTCCATGAACAG 89261  
 QY 3796 CCCTTGAAGCGCGCGAGTATGGGAAGTCTCAACAGGAGCAGGTGGCCCTCGGTGCA 3855  
 Db 89260 CCCTTTAAGACCCAGAGTATGGGAGCCAGCTCAACAGGAGCCACTGACTTCCAGTGTCT 89201  
 QY 3856 CTCGACGCTGCTGTGTGCGGGATTTCAAGCCTCAAGAGCTGAAGAGCAGCCCGCTCAAG 3915  
 Db 89200 CTAGATCATGCTGTGTACAGGATTTCAAGGTTCAGAGCTAAAGGCAACAGCTTGCAG 89141  
 QY 3916 GGGAGCGGGGCGCAGTGAATTTTCGGCTCCCGGTAGCGCCCAATGAGTCAAGTGGCAGC 3975

```

RESULT 10
AC068502
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC068502
175144 bp DNA linear HTG 03-MAY-2000
MUS musculus chromosome 13 clone RP23-305J14 strain C57BL6/J,
WORKING DRAFT SEQUENCE, 9 unordered pieces.
AC068502
AC068502.1 GI:7684387
HTG; HTGS_PHASE1; HTGS_DRAFT.
MUS musculus
MUS musculus
Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 175144)
Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L.,
Huang,M.C., Idol,J., Lee-Lin,S.-Q., Maduro,O.L., Maduro,V.B.,
Mastrian,S.D., McCloskey,J.C., Morse,E., Ojodu,M.A., Pearson,R.,
Stantripop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
Tiongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
Wetherby,K.D. and Green,E.D.
NISC Mouse Sequencing Initiative
Unpublished
2 (bases 1 to 175144)
Green,E.D.
Direct Submission
Submitted (03-MAY-2000) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@hgri.nih.gov
----- Project Information
Center project name: xe
Center clone name: 305J14
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 16396 bases at least Q40
Consensus quality: 170792 bases at least Q30
Consensus quality: 171793 bases at least Q20
Insert size: 144000; agarose-fp
Insert size: 177000; pulse-field-gel
Quality coverage: 8.31x in Q20 bases; agarose-fp
Quality coverage: 6.76x in Q20 bases; pulse-field-gel
Quality coverage: 6.83x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3763: contig of 3763 bp in length
* gap of unknown length
* 3764 12308: contig of 8545 bp in length
* gap of unknown length
* 12309 18926: contig of 6618 bp in length
* gap of unknown length
* 18927 40898: contig of 21972 bp in length
* gap of unknown length
* 40899 57495: contig of 16597 bp in length
* gap of unknown length
* 57496 82463: contig of 24968 bp in length
* gap of unknown length
* 82464 105126: contig of 22663 bp in length
* gap of unknown length

```

\* 105127 137979: contig of 32853 bp in length  
\* 137980 175144: contig of unknown length  
FEATURES  
Location/Qualifiers  
1..175144  
/organism="Mus musculus"  
/strain="C57BL6/J"  
/db\_xref="taxon:10090"  
/chromosome="13"  
/clone="Rp23-305J14"  
/clone\_lib="RPI mouse BAC library 23"  
BASE COUNT 52199 a 36056 c 36018 g 50864 t 7 others  
ORIGIN

Query Match 32.2%; Score 1625.2; DB 2; Length 175144;  
Best Local Similarity 78.7%; Pred. No. 6.4e-294;  
Matches 2025; Conservative 0; Mismatches 518; Indels 29; Gaps 6;  
QY 2485 GGCACAGTCCCAACACACCTGCGCTGAGTGGGTGGGCGCCATGACGCTTCTCCCGGGCAGA 2544  
Db 166362 GGCACAGTCCCAATAAACAATATAGCTCGGTGGGCGCGGACCTCTCCCGGACAGA 166421  
QY 2545 AGCGACTCTCTGGGTGGAGCTCACTATGCTGAACATGCTCAACAGAGGAGCAGCAGC 2604  
Db 166422 AGTGAAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 166481  
QY 2605 GCCAGCAGCATCAGCTCGGCTACCTGAGCAGCGCGCTCTCAAGGGATCTCGCCCTGC 2664  
Db 166482 ACCAGCAGTATCAGCTCTGCTTACCTGAGCAGCGCGGATCTCGGGCATCTCCCGCTGC 166541  
QY 2665 TTCTCAGCGCGCGCTCCAGCGAGGCGTCCAGCGCGCGGCGCGCGGAGTCTGAGC 2724  
Db 166542 TTTTCCAGCGCGAGTCCAGCGAGGCGTCCAGCGTGAAGGCGCGCGCGAGTGTGAGT 166601  
QY 2725 GTGGCGACTCTACAGCCCATCTCCAGCGAGCGCTCGCGCGCTCCAGCGAGGAGCAGC 2784  
Db 166602 GTGGCGACTCTACAGTCCATCTCCAGAGTCTTCAAGAGGAGTCCAGCGAGGCGCAGC 166661  
QY 2785 CAGAGCGAGCGCTGCCAGCGCTGCTCAGCTCAGCGCGCGCGCGCGCGCTCAG 2844  
Db 166662 CAGGCTGAGCGGCTGCCAGTCTGCTCAGCTCAGCGCTCAGCGCGCGCGCGCTTAC 166721  
QY 2845 GCCAATAGCGCGCTGCCAGAGAGGCGCGCGCGCGCGCGCGCTGCCAATAGAGAGG 2904  
Db 166722 GCCAATAGTGCAGTGCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 166781  
QY 2905 ATGAGCTGAAGAGCGCGCTGGCGTGTCTGGGATGCTCGCGCTGCGCGCTGGCGCTG 2964  
Db 166782 CTGAGCTGAAGAGCGCGCTGGCGTGTCTGGGATGCTCGCGCTGCGCGCTGGCGCTG 166841  
QY 2965 CTCCAGTTCATGCCCGGAGGAGTGCAGCGAGCGCGCGCGCGCGCTACGGCGCGCGC 3024  
Db 166842 CTCCAGTTCATGCCCGGAGGAGTGCAGCGAGCGCGCGCGCGCTACGGCGCGCGC 166901  
QY 3025 CAGCTGCAGCGCGAGTGCCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3084  
Db 166902 CAGCTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 166961  
QY 3085 ACAGGCTCCGAGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3144  
Db 166962 ACAGTCTCGAGAGCAATGTCATCTGCGAGGCTTCAACCGCTTCAACCGCTTCAAC 167021  
QY 3145 AACCCCGCGCGAGTGCAGCTGCCGCGGAGAGCGAGTCTGCTGCTGCTGCTGCTGCTG 3204  
Db 167022 AATCTTCAAAATTTGCTCCATCGTGGGAGAGAGTACGTTAGTGTGCTGCTGCTGCT 167081  
QY 3205 CGCGCGAGGCGCGAGTCCCGAACTTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3264  
Db 167082 CGCGAGGAGAGCGAGCAACCGCGTACTTCCAGGATCCCTTGGCGCTCCCGAGCATCA 167141  
QY 3265 GAGAGCTCAGCTGGAGTCCCTGACCATGAGCGTGTGATGCCAAGCTGAAAGATGAGAT 3324  
Db 167142 GAGATGTTGCCCTGAGGCTCTGACCATGAGTGTGATGCTGCTGCTGCTGCTGCTGCT 167201

QY 3325 TTCTCTCCGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3384  
Db 167202 CTCCTGCCAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 167261  
QY 3385 CACTTCCCGAGCGCTCTCCCGGAGCAGCAAGTGCCTCCCGGCGCGCGCTGCTGCTGCTG 3444  
Db 167262 CAGTTGCAGTGGCATCTCTGAAAGACAGCAAGTAGCCCATGAGCAGAGAGCTGGAC 167321  
QY 3445 GCGCCCGGCTGCGCAGAGCAGCAGCTGCGCAGAGTTCATGCTGCGCTGCGCAGCAGC 3504  
Db 167322 TTAGCAGGCTGCCAGACAGTCTGTTGGCAGGAGTACCAGCTGCGCAGCAGCCTGC 167381  
QY 3505 CCCGAGGAGCAGCAAAACCGCTGCTCCATTCAGTGGAGCAAGTCAAGTCCCGGAGCGCC 3564  
Db 167382 TCTGAGGAGCAATAAATGATTTGCCCATCAGTGGAGAGGTCAGTCTGGAAGTCT 167441  
QY 3565 GACCTGTCTCTCTCCAGCTCAAGTGTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 3621  
Db 167442 GATCTGTCTCTCTCCAGCTGAAGTGTGTCAGCAGCGCGCTGAGCAGCAGCAGCCTCG 167501  
QY 3622 GCTTTGGGTTCTGCAAGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3675  
Db 167502 GCTTTGGGCTATACAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 167561  
QY 3676 GGGCTGCTGGGCTATCAGACCTCGGGGAGCAACAGCAACCTAGGTTGGCGCGCAG 3735  
Db 167562 GCGCAGCTGGGAGTATCAGACCTCGGGGAGTATGAGTATCTGCTGCTGCTGCTG 167621  
QY 3736 CACTTGTATCTCCACAGCGCGCGGAGTGGCAGTGGAGCAAGCTTCCATGAACAG 3795  
Db 167622 CACTTGTCAATCCACAGTGGGATGGTGGCAGCAATGAAATACTTCCATGAACAG 167681  
QY 3796 CCCTGTAAAGCGCGCGAGTATGGGAACTGCTCAAGAGCAGCAGTGGCGCTGCTGCTG 3855  
Db 167682 CCCTTAAAGCGCGCGAGTATGGGAGCGAGTCAACAGCGCAGTCTGCTGCTGCTG 167741  
QY 3856 CTCGACGCTGCTGCTGGTGGCGGATTCAGCCTCAAGCTGAAGAGCAACCGCTGCTG 3915  
Db 167742 CTGATCATGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 167801  
QY 3916 GGGAGCGGGCGCGAGCTGAATTTCCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 3975  
Db 167802 GAGAAATGGGGTTTGGCTAGATTTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 167861  
QY 3976 ATGCTGAATGCTATGCAAGACCGAGCAGCAGTGGGAGGAGTGGCTGCTGCTGCTG 4035  
Db 167862 ACAGTGAATGCTATGCAAGACCGAGTCAATGCGGAGGAGTCAATGCGGCTGCTG 167921  
QY 4036 CTCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4095  
Db 167922 CTCAGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 167981  
QY 4096 ATTAGTGTCTCTCACATCAACATCTTCCAGGCGCAGAGAGTGGCTGCTGCTGCTGCT 4155  
Db 167982 GTTGGTGTCTCTCACATCAACATCTTCCAGGCGCAGAGAGTGGCTGCTGCTGCTG 168041  
QY 4156 CACGCGATGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4215  
Db 168042 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 168101  
QY 4216 AGCTTTGGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4275  
Db 168102 AGCTATGGGCGTACAGCGTCAAGCAATGCCAAGGGCGCAACCTCACTCTGCAACAGGA 168161  
QY 4276 CAGCTCAGTGCACAAAGTCAAGCCTGAGGCTGAATGATCAAGATGAGATGAAGAGG 4335  
Db 168162 CAGCTCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 168221  
QY 4336 CAGCCCATCGCTGCTCTAATCTCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4395  
Db 168222 CAGTCCCGAGCAGCTCTGCTTACCGTCAAGAAATATTATTCGGTCTGCTGCTGCTGCTGCT 168281

Copied from PCI US2003092 on 04-03-2004

1075	GAATATCAGCCCTGCGCTTGAGCTTCACTACTCTTCGGCGCCGTTCTCTCTC---CACATG	1133
a		
b		
721	GCATTAAGCCCTGCTGTGAGTTTCACATACCCCTCTACACCGGTATCCTCCAGCAAAAT	780
a		
b		
1132	CATCAGCAGATCCTTAAGCCGACAACACAGAGCTTAGGTTAGCCTTTGGACACAGCCCTTCCA	1191
y		
a		
b		
781	CATCAGCAAAATTAAGTCGTGACAAACCCCTGGGTCTTGCCCTTTGGACACAGCCCTCTCT	840
a		
b		
1192	CTCATCCACCTGCGCCCAACTTTTCCAAACACAGAGAGSCCTATTTCAGAGGATCCCTACGGTT	1251
y		
a		
b		
841	CTCATCTCATCTGCTCCAACTTTTCTTCCAGAGACCTATCCCTGGCATCCCCAGTGTCT	900
a		
b		
1252	CTGAACCCCGCTCCAGGTCAGCTCGGCGCCTCTTGAGTCTCTCA---CAGAAACAAGCCCAAG	1308
y		
a		
b		
901	CTGAACCCCTGTCCAGGTCAGCTCTGGACCTTCTAGTCTCCACACAGCAGCAAAACCCACACA	960
a		
b		
1309	AGTGAGTCTGCAGTGCAGCAGCACTGGTGACCCGATGCACAACAAGAGTTCACAATCAAA	1368
y		
a		
b		
961	AGTGAGTCTGCAGTGCAGCAGCACTGGGGATCCAAATGCACAACAAGCGCTCCAAAGATAAG	1020
a		
b		
1369	CCGATTAAGAGACCTCCCGACGCCAGGGGCTCGGGGACAGCGAGCAACCGCAAGGAACA	1428
y		
a		
b		
1021	CCGGATGAGGACCTCCCGACCGCGGTGCGGGAAGCGTCGAGAAACCCGAAAGGTATG	1080
a		
b		
1429	ACCCTTGTCAAGGAGGAGGGACAAAGATGAAAGCAAAACAGAGGCCTGAAGTCATCTAT	1488
y		
a		
b		
1081	ACCCCGGTAAAGGAGGAGGGACAAAGATGAAAGCAAAACAGAGSCCTGAAGTGSTCTAC	1140
a		
b		
1489	GAGACAACCTGCGACTGGGAAGGCTGGCGAGGGATTCGCACCCCAAGACAGCCTTGCTG	1548
y		
a		
b		
1141	GAGACGAACCTGCCTATTGGGAAGGCTGTTCGCCGGAGCTTTGACACACAGGAACAGCTAGTG	1200
a		
b		
1549	CACCATATAAATACGACCATATTCATGAGAGAGAAGAAGAGCTTCGCTGCAGCTGGCTG	1608
y		
a		
b		
1201	CATCATATAAATAATGACCATACATACATGGTGAGAAGAAAGATTTGTGTCCGATGGCTG	1260
a		
b		
1609	GACTGCTCAAGAGAGCAGAAACCCCTTCAAAGCCGAGTATATTTGGTAGTCGATATGAGA	1668
y		
a		
b		
1261	GACTGTTCGCCGGAGCAGAAACCATTCAAAGCCAGTATATGCTGSGTGCACATGAGG	1320
a		
b		
1669	AGACACACGGGGAGAGAGCCTCACAATGCACTTTTGAAGGTTGCACAAGAGCCTACTCG	1728
y		
a		
b		
1321	AGACATACAGGGGAAAAGCCACCAATGCACTTTTGAAGGTTGTACAAGAGCCTACTCC	1380
a		
b		
1729	AGACTAGAAAACTTTGAAAAACACTTTGAGATCTCACACTGGAGAGAAAACCATAGCTGTG	1788
y		
a		
b		
1381	AGACTAGAAAACTTTGAAAAACACTTTGAGATCTCACACTGGAGAGAAAACCATATGTGTGC	1440
a		
b		
1789	GAGCACGAAGGTTGCACAAGGCTTTCTCAAATGCCTCTGATCGCGCCCAACACACAAAAC	1848
y		
a		
b		
1441	GNACACGAGGCTGCACAAAGCTTCTCCAATGCACTGTGACAGGGCCACAGCACCAAAAC	1500
a		
b		
1849	AGAAGCGCATTCOATGAGAAACCATATGTGTGCAAAATCCCAGGCTGCACCTAAGCGTTAC	1908
y		
a		
b		
1501	AGGACTCATTAATGAAAAACCATATGTTTGAAGATCCCAGGCTGCACAAAAGCGCTAT	1560
a		
b		
1909	ACAGACCCCAAGCTCCCTCCGGAACATGTGAAGACAGTGCATGGCCCCAGAGAGCTCATGTC	1968
y		
a		
b		
1561	ACAGACCCCTAGTTCTTCCGGAACATGTGAAGACTGTGCATGTGCCCTCAGAAACACAGCTC	1620
a		
b		
1969	ACCAAGAGCAGCAGGGGACATCATCTCTGCGCGCCACCCCGCAGAGATTTCGGGAGC	2028
y		
a		
b		
1621	ACCAAGAGCAGCGAGGAGACATCCACCNAAGGCTTCCACACCAAGACAGCCAGGCAGC	1680
a		
b		
2029	CATTTCAGAGTCCAGSTCGCCTGGCCGACCGACTCAGGAGSCCTTTGGTGACACAGCAGAC	2088
y		
a		
b		
1681	CACCTCTCAGACAGCATCACAGGCCACAGACTCAGGGTGCAGCTGGGAGCAGAAAGAC	1740
a		
b		
2089	CTCAGCACACTACCTCNAAGCCGGGAAGATGCTCCAGGTGAAACCGTCAAGGCGAG	2148
y		
a		
b		
1741	CTCAACAGCATCTACTCAAGCGGTGAAGATGCTTCCAAGTGAAGAGCACTCAAGTCTAGAG	1800
a		
b</		



Copied from PCI US2003092 on 04-03-2004

[illegible]



Db 107581 CTCATTCTCCCTCTCAGAGGCCAACTGCTGCTCAGGAGATGGCTCTGAAATTCATGA 107522  
 QY 4488 GTTACTTTCCCGAGGTGCTTAATCAGGTGACAAGCACAGTGGACACCTTCGACAGCCATGA 4547  
 Db 107521 GTTACTTTCCCGAGGTGTTAAACAGGTGACACAGTGGACACCTTTGAGAGTCAATGA 107462  
 QY 4548 CTTGGAAGGGTACAGATTGACTTCGATGCCATCATAGCATGGGACCACTCCAGGCT 4607  
 Db 107461 CTTGAAGGCGTGCAGATTGATTTTGTATGCCATCATAGATGGGACCACTACAGGCT 107402  
 QY 4608 GATGTCCGGGGCCCTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGGCT 4667  
 Db 107401 AATGTACAGGGCCTTGAGCCCAAGTATCATTCAGAACCTTTCCCATAGCTCTCCCGGCT 107342  
 QY 4668 CACAGGCTCCGGGGCTCCCTCCCATTCAGTCCAGTCCGCTGTCATGAGCAGCACCACATGG 4727  
 Db 107341 CACACACCTCGGGCATCCCTCCCATTCCTCCTGTCATGAGCAGCACCACATGG 107283  
 QY 4728 CTATCGGGGACATGAGTCTTTGCTGACCTCCCTAGCGGGAAGAAAGCAATTCCTTTCGAG 4787  
 Db 107282 CTATGGGATGATGAGTCTTTGCTGACCTCCCTGCGAGCAAGAAAGCAATTCCTTTCGAG 107223  
 QY 4788 TTATGCAATAGGCTTTAGGAAAAAAGACTGCAACCAACGAGAA 4831  
 Db 107222 TTATGCAATAGGCTTTAGGAAAAAAGACTGCAACCAACGAGAA 107179

## RESULT 13

MMGLI2  
 LOCUS M.musculus mRNA for gli2 gene. 5849 bp mRNA linear ROD 06-FEB-1997  
 DEFINITION X99104  
 ACCESSION X99104.1 GI:1707589  
 VERSION gli2 gene.  
 KEYWORDS Mus musculus.  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 5849)  
 AUTHORS Hughes,D.C., Allen,J., Morley,G., Sutherland,K., Ahmed,W., Prosser,J., Lettice,L., Allen,G., Mattei,M.G., Farrall,M. and Hill,R.E.  
 TITLE Cloning and sequencing of the mouse Gli2 gene: localization to the dominant hemimelia critical region  
 JOURNAL Genomics 39 (2), 205-215 (1997)  
 MEDLINE 907179222  
 PUBMED 9027508  
 REFERENCE 2 (bases 1 to 5849)  
 AUTHORS Hughes,D.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-JUL-1996) D.C. Hughes, University of Nottingham, MRC Institute of Hearing Research, University Park, Nottingham, Notts, NG7 2RD, UK

## FEATURES

Source  
 1..5849  
 /organism="Mus musculus"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 1..5849  
 /gene="gli2"  
 <1..>5849  
 /gene="gli2"  
 1334 a 1876 c 1524 g 1115 t  
 BASE COUNT  
 ORIGIN

Query Match 17.1%; Score 864.8; DB 10; Length 5849;  
 Best Local Similarity 60.3%; Pred. No. 1.4e-151;  
 Matches 1836; Conservative 0; Mismatches 1067; Indels 141; Gaps 19;

QY 413 CCCACTACCACTCCTCCTCTTTCCCTGCTTCATCCTCTGACCAATGATGCA 472  
 Db 237 CCCAGGAGTTCCTCAGCAGCTCCTCGCGCTTTCCACGCGCTTTGCGGATGACATGA 296

QY 473 GACATCATGAGGCGCTTACCATTAACGATCAATCCGATTCCTCAATTCATATGACTT 532  
 Db 297 GACACAGGAGGAGGTACCATTTATGACCTCACTCTGTCACAGTGTACACGGGCTC 356  
 QY 533 CCGCTTATCTAGTAGCCCTACGCTATCCGGACCTTCGCTTCATTTAGGATCTCCCAACCC 592  
 Db 357 CCACCTTAAGTGGCAGCCCTGTCATCTCAGATATCTCTTGTATGACGACTTCTCCAC- 413  
 QY 593 GGAACCCGCTGCTGCTTCGAGTCTCCCTTCAGCCCTCCACATCCCATTAATTCCT 652  
 Db 414 ---ACCCTGCTGGCCCTGGAGGTCACCCCTTCAGCGCCACACCCCTAGCTGAACCC 470  
 QY 653 ACATGG---ACTATATCCGCTTCGACAGCAGCCCATCGCTCTCCATGATCTCAGCAA 709  
 Db 471 ATATGGAGCACTACTCTCGGTCTGTGCACAGCAGCCCACTCTCAATGATCTCTGCGG 530  
 QY 710 CCCGTGGGCTGAGCCCTACAGATCGGCCCATG----- 742  
 Db 531 CCAGGGGCTCAGCCCTGCTGATGTGGCCCAAGAACTCTGAAAGAGAGGGGACTCTTTA 590  
 QY 743 -----CAGGAGTCAGCCAGCAGATACTATCATCATGAGCCCTGCTACTG 790  
 Db 591 GCCTGGAGCCCGCAGGACCAACCTTCAGACTATTACACAGATGACCCCTCATGGCAA 650  
 QY 791 GCCAGCGAGCCCTATGACAGACATTTATCCCTCAGCTGCCACCCGCGCAGCGGGCCA 850  
 Db 651 GCCACCCACCCCTTATGGGACCTTCTAATGACAGAGGGGGTCTGCTAGCGCACCCC 710  
 QY 851 TCCACATGGAATATCTTCACTATGATAGCAGCAGATTTCCAGGCCAGGCTGTGAG 910  
 Db 711 ATCTCCATGACTACTCTCAACCTGTGGATGATCAGGATTTCTTAGTCCAGCTGTGACC 770  
 QY 911 CCAGCCGAGCGGAAAGCTGACACTGTCATATCACCACCTCTCCGATCATAGCTTGACC 970  
 Db 771 CACGACTCAGCGCAAGCGGCTCTGTCATCTCCCGCTCTCAGATGCCAGCTCGACC 830  
 QY 971 TTCAGACCATGATAGGAGCTCTCCAACTCTTGGTCACGATTTCTAATTAATTCGGTA 1030  
 Db 831 TACACGATGATTTGGGACCTCTCCCACTCGCTGGTAGCTTACATCAACACTTCCAGGA 890  
 QY 1031 CGAGCTTTTCAGCAAGTGGCTCTTATGCTACTTATCTGCAAGTGCATACGCTGCT 1090  
 Db 891 CGAGCTCAGCAGCAGTGGCTCTTATGGACATCTGCTGCTGGTGGCTCAGCGCAGCT 950  
 QY 1091 TGAGCTTTCACCTACTCTTCCGGGCGCTCTCTCCACATGATGATGATGATGATGAT 1150  
 Db 951 TCACCTTTCCCGC-----CCCATCAATCTGTGGCTACACGAGATCTTGAGCC 1001  
 QY 1151 GACACAGAGCTTAGGTTTCAGCTTTGGACACAGCCCTTCCACTCATCCACCTGCCCAA 1210  
 Db 1002 ACCAGCGGGCTGGGCTCAGCTTTGGACACACACACACCTGATCAGGCTTCCACCA 1061  
 QY 1211 CTTTTCACACACAGAGGCTATTCAGGATCCCTTACGGTTCCTGAAACCCCTCCAGGTCA 1270  
 Db 1062 CTTTTCAGCAGCAGCAGCTGCTCAGCTCCATCAGCAGCAGCTGCTCAGCAACTCA 1121  
 QY 1271 GCTCGGC-----CCTTCTGAGTCTCAGACAAAGCCCGAGTGTGCTG 1318  
 Db 1122 GCAGCAGTAGCAGCAACTGTCTAAATGATGCCAACCAAGCAGAACGAGCTGAG 1181  
 QY 1319 CAGTCAGCAGCAGCTGGTGACCCGATGCACA---ACAAGAGTTCAGATCAAAACCCGATG 1375  
 Db 1182 CTGTGAGCAGCAGCTGAGCCCATCACCATTCAATGAGGAGCAAGTCAAGACTGAGG 1241  
 QY 1376 AAGACCTCCCGAGCCCGAGGGGCTCGGGGCGCAGAGGACAGCCCGAGGACAAACCTTTG 1435  
 Db 1242 CTGAGGGCTGCTCCAGCATCCCGCTTGGACTGCACAG---GAGCAGCTGGCCGATC 1298  
 QY 1436 TCAAGGAGGAGGGGACAAAGATGAAGCAACAGGAGCCCTGA---AGTCATCTATGAGA 1492  
 Db 1299 TCAAGGAGAGCTGAGGAGGATGACTGTAAGCAGAGGGCGGAGGTGTCTATCAGGAG 1358  
 QY 1493 CAAACTGCCACTGGGAAGGCTGCGCGAGGAGTTTCGACACCCCAAGAGCAGCTTGTGTCACC 1552

1359 CCAACTGCCACTGGGAGAGTGCACCAAGGAGTATCACACACAGGAGCAGCTGGTGCATC 1418  
1553 ATATAATACGACCATATTCATGAGAGAGAGAGAGTTCGTGTGCAGGTGGCTGGACT 1612  
1419 ATATCAATATGACATATCCACGGGAGAGAGAGTTCGTGTGCAGGTGGCTGGACT 1478  
1613 GCTCAAGAGAGAGAAACCTTCAAGGCCAGTATATGTTGGTAGTGCATATGAGAAGAC 1672  
1479 GCAGGAGAGAGAGAGAGCTTCAAGGCCAGTATATGTTGGTAGTGCATATGAGAAGAC 1538  
1673 ACAGGGGAGAGAGCTTCAAAATGACATTTTGAAGGTTGCAAAAGGCTTACTCGAGAC 1732  
1539 ACAGGGGAGAGAGAGCTTCAAAATGACATTTTGAAGGTTGCAAAAGGCTTACTCGAG 1598  
1733 TAGAAAATTTGAAACACACTTCAGATCTCAGATCTCAGATCTCAGATCTCAGATCTCAG 1792  
1599 TGAGAACCTTGAAGACACACTTCGTTTCAACACAGAGAGAGAGCTATGTTGTGAAC 1658  
1793 ACAGAGGTTGCAAGAGGCTTTCTCAATGCTCTGATCGCGCCAAACACCAAAACAGAA 1852  
1659 ACAGAGGCTTGAACAGGCTTCTCAATGCTCTGATCGCGCCAAACACCAAAACAGAA 1718  
1853 CGCATTTCAATGAGAAACCATATGTTGCAAAATGCTGCAAAAGGCTTACTCGAGAC 1912  
1719 CTCACCTCCAATGAGAAACCTTCAATGCTCTGATCGCGCCAAACACCAAAACAGAA 1778  
1913 ACCCAAGCTTCCGGAACATGTGAACAGAGTGTGATCGCGCCAAACACCAAAACAGAA 1972  
1779 ACCCAAGCTTCCGGAACATGTGAACAGAGTGTGATCGCGCCAAACACCAAAACAGAA 1838  
1973 AGAAGCAGGAGGAGACATTCATCTCGCGCCAAACACCAAAACAGAA 2032  
1839 AGAAGCAGGAGGAGACATTCATCTCGCGCCAAACACCAAAACAGAA 1898  
2033 CACAGTCCAGGTCGCTTGGCCGACCATCTCAGGAGGCTTGGTGTGAGAGAGGACTTCA 2092  
1899 AGCCAGGCGCGGAGGAGTGGCC-----GGGACCTGAGGAGAGTGTGGAGGCA 1949  
2093 GCAACACTACCTCAAGCGGAGAGATGCTCCAGGTGAAACCGTCAAGGAGAGAGAGC 2152  
1950 GTAGCAGGAGGAGGAGTGGAGGAGTGTGATCAATCAAGGAGAGAGAGC 2009  
2153 CAATGAGATCTCAGCAAGGCTTGGTGTGAGTGTGATCAAGGAGAGAGAGC 2212  
2010 CCGGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2069  
2213 TCAGCAACTATTCCA-----ACAGTGGGCTGAGCTTCTCTGACGATGAGTAGTA 2266  
2070 TGGGAGTGGGCGGCGGAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2129  
2267 TAGGAGACTCAGTGTGATGATGAAC-----CCCAATCATGAGTCAACATTTCCACTG 2323  
2130 TGGGAGACTCAGGAGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2189  
2324 CACACAGGCTTGTGTTGCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2383  
2190 CACCTTCCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2249  
2384 ACCTAACTAGAGAGGCTTAAACAGAGTGAATGGAATGTTTCCGAGCTGAGACCCATTC 2443  
2250 AGCTGAAGAGAGAGAGAGTCAAGTCACTGAAGGATTCCTGCTGTGCTGTGCTGTGCT 2309  
2444 TACCCCTTAAAGCCCTGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2497  
2310 CACACAGGCTTAAACAGAGTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 2369  
2498 ACAACAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2557  
2370 TCAATATACAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2429  
2558 GGTGAGCTCACTATGCTGA-----ACATGCTCAACAGAGAGAGAGAGAGAGAGAG 2614

2430 TCACCAAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2489  
2615 TCAGTTCGGCTTACCTGAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2674  
2490 TGAGTTCGGCTTACCTGAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2549  
2675 GCGCTTCAGCAGGAGTGTGAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2731  
2550 GTCGCTTCAGCAGGAGTGTGAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2609  
2732 ACTCCTACGACCGCTTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2788  
2610 ACTCCTATGACCGCTTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2669  
2789 GCGAGGCTTCTCAGCAGCGCGCTCTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCT 2848  
2670 GTGGGCTTCTCAGCAGCGCGCTCTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCT 2729  
2849 AGTACGCTTCTCAGCAGCGCGCTCTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCT 2908  
2730 AGTACGCTTCTCAGCAGCGCGCTCTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCT 2789  
2909 GCGTGAAGCAGCGCTTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2965  
2790 GCGTTCGATACCGCTTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 2849  
2966 CTCAGTTCATGCGCGAGGAGTGTGAGCAGCGCGCGCTCTCTCAGGAGTGTGCGCTTCT 3025  
2850 CACATTCATGCGCGAGGAGTGTGAGCAGCGCGCGCTCTCTCAGGAGTGTGCGCTTCT 2909  
3026 ACCTGCA-----GCGCAGCAGTGTGCGCGCGCTCTCTCAGGAGTGTGCGCTTCT 3064  
2910 ATGCTACGAGTGTGCGCTTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCT 2969  
3065 GGGCAGGAGCGCTTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 3124  
2970 GGGCAGGAGCGCTTCTCAGCAGCGCGCTCTCTCAGGAGTGTGCGCTTCTTCCAGCC 3026  
3125 TCAGCAGCTCAGCAGTGTCAACCGCGCGAGTGTGAGCAGCGCGCTCTCTCAGGAGTGT 3184  
3027 TCCACAGTACCCACACATGATCCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 3086  
3185 TCGTGTTCAGATTTACACGCGCGCGCTCTCAGGAGTGTGAGCAGCGCGCTCTCTCAGTTC 3244  
3087 TGCAGTACAGAGCGCGCGCTCTCAGGAGTGTGAGCAGCGCGCTCTCTCAGTTCAGTTC 3143  
3245 CCGTTCCTCCAGATCACCGAGAGTGTGAGCAGCGCGCTCTCTCAGGAGTGTGAGCAGTGT 3304  
3144 CCAGACCGCTTGTGATGAGAGAGTGTGAGCAGCGCGCTCTCTCAGGAGTGTGAGCAGTGT 3203  
3305 CCAACCTGAGAGTGTGAGTGTGAGCAGCGCGCTCTCTCAGGAGTGTGAGCAGTGTGAGC 3348  
3204 GCCAGGCTGAGTGTGAGTGTGAGCAGCGCGCTCTCTCAGGAGTGTGAGCAGTGTGAGC 3247

## RESULT 14

AF109923

LOCUS

DEFINITION

Xenopus laevis zinc finger protein G112 (G112) mRNA, complete cds.

AF109923

VERSION

AF109923.1

KEYWORDS

SOURCE

Xenopus laevis.

Xenopus laevis

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopus laevis.

Xenopus laevis

REFERENCE

1 (bases 1 to 4517)

AUTHORS

Lee, J., Platt, K.A., Censullo, P. and Ruiz i Altaba, A.

TITLE

G112 is a target of Sonic hedgehog that induces ventral neural tube

JOURNAL

development 124 (13), 2537-2552 (1997)

MEDLINE

97359968

```

9216996
REFERENCE
2 (bases 1 to 4517)
Brewster, R., Lee, J. and Ruiz i Altaba, A.
Gli2/zic factors pattern the neural plate by defining domains of
cell differentiation
JOURNAL
MEDLINE
Nature 393 (6685), 579-583 (1998)
9634234
REFERENCE
3 (bases 1 to 4517)
Brewster, R. and Ruiz i Altaba, A.
Direct Submission
TITLE
Submitted (28-NOV-1998) Cell Biology, Skirball Institute, NYU
Medical Center, 540 First Avenue, New York, NY 10016, USA
FEATURES
Location/Qualifiers
1..4517
/organism="Xenopus laevis"
/db_xref="taxon:8355"
1..4517
/gene="Gli2"
296..4360
/gene="Gli2"
/Note="similar to Mus musculus Gli2 and Xenopus laevis
Xgl14"
/codon_start=1
/product="zinc finger protein Gli2"
/protein_id="A028180.1"
/db_xref="GI:4704617"
translation="MEHYLRVSHNSPTLSMISAAGLSPAEVAHEHLKERGIFGLAPP
PGTPTTTCYCHQMAFLASHPYGLLELVQSAAGANTSHLDYLTMDVSRFSSPRVTP
RLSRKRLASIPDSADLODMITRSPNLSVAYINNSRSSASAGSYGHLSSAGAIAP
AFSPHPNPVAYOCLLSQOSSLSSFGHTPLLPSPFASRQOGLVLSANPAPSNNN
NPDSILNKVESASVSYNOVIHKKRWKTEEEVLRDSSPDHLTDLEDKDDKE
KQEPDIIYENCHWDGCKEFQDQLVHHINDDHIEGKEFEVCRKESQKQF
AFQYMLVHMRHRTGKPHKTEFCFKRAYSRLENLKLHRSHTGKPYVCEHGGCKR
KAFNSADRAKQNRTHSNEKPYIKIPGCTRYTDPSSLRKHVKTGHGPEAHVTKYR
NDIIQKSPVXENGNEASKLSSREHSSTYRDSQCLQARTIKTEHNMHQSPPG
QSSCSSEPSYSNYIDSGVDVLSHGEGLGLLEETNPVVDVSSASPAVGLQLR
KHGTYPNRDLQKLEKLTMKESCFINPPPARNTPLITANGVSAIGVBSVLE
NPRATELPMNDTLNLQNERDSTSSLSAYSRSSGSLSPFSRSETSQFGG
RUNSSADSDYPTISDASRRSSSEASHGLPNLLNLPQOYRLKARYAAATGGPPP
TPLANMRINTLNKLSMDRPFPLHAFRQLPVRPRCSGGGNAGLAVYPIEFGNNS
RASDPVARTGMDQGLPVRQRFHSMNMTLHPSLSERRNGLOHYNRSDSLHR
HVSPPRPSINENVAEMAIISGLDVPDGGDDMLPDDVVQVIRSONREAPKNLPIEY
NSPARNLOSSIKPFPHHTPHRAPGAYSRRNPAEAECILDOTANMODNSMPYQWNEVS
GTVDYSLDKPQFPGNLAIVQOQKFAOYQNFNOAPIHGAQNTINGQCESYQNRISV
NGQRENTWQOQMQSCQNVTSDFSPQRYSQSOMLNRAMQEGNLYLPPSCNNMA
ERPGVHTHSASNTLSHQHLAVHGAPTKMTNFNSVNDQGLHPNAYTVQKNGLEP
QOATLGMQOAFNHGMIQPRPAAPHLNPRNTHPVHPPYMRSPVLSLSPQQT
AEATPKRTNEDTPSPKDNLLYYSGQIHMYEPNGNFSGIDCTAQOQLTPWSPGANO
VTSTVDSQGLEHPVVIDFDTMMDGDHSLMSGLTSLPGLLHFSQTSRLTTPNSLT
LPSIPAGISNMAIGDMSSMLTTLABESFELNLS"
BASE COUNT 1295 a 1292 c 990 g 940 t
ORIGIN
Query Match 15.8%; Score 796.2; DB 5; Length 4517;
Best Local Similarity 58.2%; Pred. No. 9, 5e-139;
Matches 1761; Conservative 0; Mismatches 1128; Indels 138; Gaps 15;
QY 428 CTCATCTTTTCCGCTCCATCCCTCCTGATACCAATTGATGCGACATGAGGGCC 487
D 11111111111111111111111111111111111111111111111111111
D 81 CTCCTCTCTCCGCTCCACCCCTCCCTTCCCAATTGAAAGTGGCCCAAGAGGAA 140
QY 488 GTTACCATTCAGATTCATCCGATTCCTCCATTCGATGATGACTTCGCGCTTATCTAGTA 547
D 11111111111111111111111111111111111111111111111111111
D 141 GATACCATTCAGACATTCATCCATTCATGAGCTCTCCCTCTCACTCTTGGGTACACAG 200
QY 548 GCCTTACGATCCGAGCTGCCCTTCATAGGATCTCCCAACACCGGACCCCGCTGCTG 607
D 11111111111111111111111111111111111111111111111111111
D 201 GCCTGTGATTCAGACATCTCCTTAATAGCGCTCTCCCTCTCACTCTTGGGTACACAG 260
QY 608 CTTCCGAGTCTCCCTTCAGCCCTCCACATCCCTACATTAATCCCTACATGAGTATATCC 667
D 11111111111111111111111111111111111111111111111111111
D 261 AGTTTGCCCATGCCACCACTATGTGAGCCGCCACATGAGGATTTACCTG-----C 311

```

1703 TTGAAGTTGCAACAAGGCTACTCGAGACTAGAAAACCTTGAACACACACTTTGAGATCTC 1762  
1365 TTTCAAGTTTTCCTTCAAGGCTTATTCAGGCTGGAGAATCTGAAGACACACTTTGCGTTTAC 1424  
1763 ACACCTGGAGAAACCATACGTTCTGTGAGCAGCAAGGTTTGCACACAGGCTTTCTCAATG 1822  
1425 ACACCGGGAGAACCGTACGTTCTGTGAGCAGCAAGGCTTGAATAGGCGTTTTCATG 1484  
1823 CTTCTGATCGCGCAACACCAACCAACGATTCCTCAATGAGAACCATATGTGTGCA 1882  
1485 CATCTGATCGCGCAACACCATCAACACAGACCCATTCCTCAATGAGAACCCCTACATCTGTA 1544  
1883 AATCCAGGCTGCTACATGAGGTTTACACAGACCCCAAGCTCCCTCCGGAAACATGTGAAGA 1942  
1545 AGATCCAGGCTGACAAAGCGTTACACGAGCCCGAGCTCTCTCAGGAACACGTCARAA 1604  
1943 CAGTGCATGCGCCAGGCTGATCTGACCAAGAACGAGGAGGACATCCATCTCTCGG 2002  
1605 CTGATGATGACCCAGGCGCATGTGACCAAGAACGAGTACCGCAATGACATTTATCCAAAAC 1664  
2003 CGCCACCCCGGAGAGATTCCGGCAGCATTCACAGTCCAGGTCGCTGCGCCACCGAGCTC 2062  
1665 CTCTGTACCAAGAGATGGGACATGAG-----CA 1700  
2063 AGGAGCCCTTGTGAGCAGGAGCTTCAGCAACACTACCTCAAGCGGGAGAAATGCC 2122  
1701 GTCCAAACTGAGCAGTAGGAGCATTCAGCAGCAGCTATAGAGACTCGGAGGACTGCT 1760  
2123 TCCAGGTGAACCGTCAAGGAGAGAACCCATGACATCTCAGCAAGCCCTGTGTGTC 2182  
1761 TGCAGGCGAGAACCTCAAAACAGAGATTAACATGATGATCAGTCCAGCCCTGGAGGCG 1820  
2183 AGTCTTCATGACAGCAACAGTCCCGCATCAGCAACATTTCCAACTGAGTGGCTCGAGC 2242  
1821 AGTCGCTTGCAGCATGAGCTTCCCATATAGCACTATATTGACAGTGAGTAGTG 1880  
2243 TCCCTCTGACCGATGAGGAGTAGTATAGGAGACCTCAGTGCCCATGATGAAC-----CCAA 2299  
1881 TGAGCCTGATGCGAGGAGGAGCTTGGGAGACCTGCTGTTGGTGGAGAGACCAATCCAG 1940  
2300 TCATGACTCAACCATTTCCATGCAACACAGCCCTTCTTGTGAAGCCGAGAAAC 2359  
1941 TGTAGACTCCACAGTGTCTATGTCAGCCCGAGCGGTGCGGCTTCAACTCCGGAACATG 2000  
2360 CGCAGGAGAC-----CAATGGATGGAGCAGTAAACACTAGAAAGCTTAAACAAAGTGAATG 2416  
2001 GTGGAACACAGTGCAGAGACTGGRACACTGAAGAGAGAAATTAAGACCATGAAG 2060  
2417 GAATGTTCCGGAGTGAACCCCATTTACCCCTTAAAGCCCTGCGGTCTCTCTCA 2476  
2061 AGTCTGCTGCTTTATTAACCCACCTCCACCTGCGAGGAACACCCGGCTCCCAACCATCA 2120  
2477 TAGGAATGSCACAGTCCAAACACACTGTCAGCTTGGTGGGCCCATGAGCTTCTCC 2536  
2121 CTGCAATGGTGTAGTGAATAGGTGTTCCAGCAGTG-----TGCTATTTCAC 2171  
2537 CGGCGAAGAGCAGCTCTCTGGGGTGGAGTCACTATGTGAACATGCTCAACAGAGGG 2596  
2172 CCAGGCGACAGACTTCCATGAATGATGTGACTTTGCTTAACCACTGAACAGAGCTC 2231  
2597 ACAGCAGCCGACAGCATCAGCTCGGCTTACCTGAGCAGCCGCGCTCCTCAGGATCT 2656  
2232 GTGACAGCAGCAGCAGCAGCTCAGTCAACAGTCCAGTGGAGGCTCCCGCATCT 2291  
2657 CGCCCTGCTTCCAGCCGCGCTCCAGGAGGCTCAGCAGCCGAGGCGCGCGCAG 2716  
2292 CTCCCTGCTTCCAGTCCAGCTCAAGTGAACCTCAGAGTTTGGAGGCGCGCTTAATA 2351  
2717 AGGTGAGGCTGCGGAGCTCTTACAGCCCATCTCCACAGCCCTCGCGGCTCCAGCG 2776  
2352 ATTCCAGCTCTGAGATCTTATGATCCCATTTCTACTGATGCGCTTAGGCGCTTAGTG 2411  
2777\*AGGCCAGCAGAGCGAGCGGCTTSCCAGCGCTGCTCAGCCTCAGCGCCCGCCAGCAGTACC 2836

2412 AGGCCAGCATCACAGTGGACTTCTTAATCTTCTCAACTCACTCCAGCTCAGCAGTACC 2471  
2837 GCCTCAAGGCCAAGTACCGGCTGCCACAGAGGGCGGCCCGCCAGCCCTGCCCCAACA 2896  
2472 GATTAAAGGCCAAGTATGCTGCAGCAACCGGTGGCCCGCCAGCCACTCTCTTCCCAACA 2531  
2897 TGGAGAGATGAGCTGAAGACGCGCTGCGCTGCTCGGGGATGCCCTCAGCCTGCGG 2956  
2532 TGGATAGGATAAATTTGAGAAATAGTTGCTCCCTGATGGACAGACAGACTTCTCTTGC 2591  
2957 TGGCCCTGCTCCATTCATCCCGGAGGAGTGCAGCGACGGGGAGCCCGACCGCTACG 3016  
2592 ATGCTTTCCGACAGCTTCTGTTCCAAAGACGTTGCAGTGATGTTG-----TGGAATG 2645  
3017 GCGGCGCCACCTCGACCGCGACGATGGCTGGCCAGCGGCTGAGAGGGCCAGCGAC 3076  
2646 CAGGACTAGCAGTGTACCCCATGAAATTCAGGGAATAATTCAGAGCAGCTAGTGACC 2705  
3077 CGGTGCGGA-----CAGGCTCCGAGGCGCTGGCCCTGCTCGCTGCTCGCGCTTCAGCA 3130  
2706 CAGTCAGAGAACTGGAGGTATGATGGTCAAGACTCCCAAGGTTTCAGCGATTTCA 2765  
3131 GCCTCAGCAGTGCACACCCCGCGGATGGCCACGTCCCGGAGAGCGCAGTCTCGTG 3190  
2766 GCATGAATAGTATGAATACATTGCACCATCCATCCCTATCGAAAGAGAAATGGAGCC 2825  
3191 TTCAGATTAACCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3250  
2826 TGCAGCACTACAATCGCTCTGACGGAAGCTTTCACAGGCAAGCT---CTATTCCTCCG 2882  
3251 CTCGCCAGTCAACGAGAACTGCTGAGTGGAGTCCCTGACCATGAGCGCTGATGCCAA-- 3308  
2883 CACCCAGCATCAATGAAATGCTCCTATGGAGGCCATATCTTGATGATTTAGATGCTCG 2942  
3309 -CTCAAGATGAGGATTTCTCGCGGAGCAGCTGGTGCAGTATTTAAATTTCCCAAGCC 3367  
2943 CGGAGAGCAGTACCTCATGCTCCAGATGATGTTGTTAGTATACATCAGATCTCAGAACA 3002  
3368 AAGCAGGCTAGCAGCAGCTTCCCA 3394  
3003 GGGAGGCACCAAGCAAAACCTACCA 3029

RESULT 15  
XLU42462  
LOCUS XLU42462 4641 bp mRNA linear VRT 26-APR-1996  
DEFINITION Xenopus laevis neural specific DNA binding protein (Xgl14) mRNA,  
complete cds.  
ACCESSION U42462  
VERSION U42462.1 GI:1150837  
KEYWORDS  
SOURCE Xenopus laevis.  
ORGANISM Xenopus laevis

REFERENCE  
AUTHORS  
1 (bases 1 to 4641)  
Marine, J.-C., Bellefroid, E.J., Pendeville, H., Martial, J.A. and  
Pieler, T.

TITLE  
An essential role for Xenopus Gli-type zinc finger proteins in the  
early embryonic patterning of mesoderm and neuroectoderm

JOURNAL  
Cell (1996) In press  
REFERENCE  
AUTHORS  
2 (bases 1 to 4641)  
Marine, J.-C., Bellefroid, E.J., Samain, H.P., Martial, J.A. and  
Pieler, T.

TITLE  
Direct Submission  
Submitted (07-DEC-1995), Jean-Christophe Marine, Laboratory of  
Genetic engineering and Molecular Biology, University of Liege,  
Allee du 6 Aout, Batiment B6-Chimie, Sart Tilman, Liege 4000,  
Belgium

FEATURES  
source  
1. 4641  
Location/Qualifiers

Query Match	15.6%	Score	789.2	DB	5	Length	4641
Best Local Similarity	58.4%	Pred. No.	1.9e-137				
Matches	1773	Conservative	0	Mismatches	1118	Indels	143
QY	428	CTCATCTTTTCCCTGCCTTCCATCTCTCTGTACCATTGATGCCACACATCATGAGGGCC	487				
Db	78	CTCTCTCTTCCCTGCCTTCCACCCCTCCCTTCCCATTTGAAATGCGCCCAACAGAGGAA	137				
QY	488	GTATACATTACGATCCATCTCCGATTTCCTCCATTGCAATGACTTCCGCTTATCTTAGTA	547				
Db	138	GATACATTATTGACACACATGGGATTCACACTTCGATGGCCCCCGACTCTCAGTGGCA	197				
QY	548	GGCCTACGATCCGGACCTTCGCCTTCATTAGGATCTCCCCACACCGGAACCCGCTGCTG	607				
Db	198	GGCCTGTGATTCAGACATTTCTTTAATCGGCTCTCCCCACACTCTCTGGGTACAACAG	257				
QY	608	CTTCCGAGTCTCCCTTTCAGCCCTCCACATCCCTCATTAATCCCTACATGGACTATATCC	667				
Db	258	AGTTTGGTCATGCCATCATATTATGTAGCCCCACATGGAGCATTTACCTG-----C	308				
QY	668	GCTCTTGACAGCAGCCCCATCGCTCTCCATGATCTCAGCAACCCCTGGGCTCAGCCCTA	727				
Db	309	GTCTGTGACACACAGCCCCACCCCTGCTATGATCTCTGAGCCCCGTGCTCAGCCCCAG	368				
QY	728	CAGATCGCGCCCATG-----	742				
Db	369	CAGAAGTTGCCCATGAGCACTTAAGAAGGGGTATCTATGGTTAGCACCTCCCGCCG	428				
QY	743	---CAGGAGTTCAGCCCCAGAGAATACTATCATCAGATGGCCCTGCTTAAGTGGCCAGCGCA	799				

QY	1880	GC	AAATCCCAGGCTGC	ACTA	AGCTTAC	ACAGACCC	AAAGCTCCCTCCG	GAACATGTGA	1939	
Db	1545	GT	AGGTC	CCAGGCTGC	ACAAAACGTTAC	ACGAGCC	CCAGCTCTCTCAG	AAACACGTTA	1604	
QY	1940	AG	CAGTGCAT	GGCCAGAGGCTCAT	GTGCACCA	GAAGCAGGAGG	GACATCCATCCTC	1999		
Db	1605	AA	CTGTACAT	GACCTGAG	GGCCATGTGACCA	GAGACCGCAATG	ACATCATC	CAAA	1664	
QY	2000	GG	CGCCACCC	CGAGAGATTC	CGCGACGC	CAATTCACAGT	CCGCTGG	CCGACCGA	2059	
Db	1665	AA	CTTCAC	TACCAAGG	AAAGCGG	GAATGAA	-----G	1700		
QY	2060	CT	CAGGAGCCCTTGT	GAGCAGCAGG	ACCTCAGCA	ACACTTACCTCA	AAAGGGGA	GAAT	2119	
Db	1701	CT	AGTGCCAA	CACTCAGCGT	AGGAGCACTCCG	CAGCTGTCTCTAG	AGACGAGC	ACT	1760	
QY	2120	GC	CTCAGGTG	AAACCGCTCA	AGCAGAGAGC	CAATGACAT	CTCAGC	CAAGCCCTGGTG	2179	
Db	1761	GC	TTGCAG	ACCAGAAC	CCATAA	ACAGAGATAC	ATGATGCAT	CTCCAGCCCGGTG	1820	
QY	2180	GT	CAGTCTT	TCATG	CAGCAGC	CAACAGTCCCC-----CAT	CAGCAACTAT	TTCAAACAGTG	2233	
Db	1821	GG	CAGTCGT	CATCAG	CAGTGTAG	CCCTCCCATATGG	CAACCAACA	CAATTGACAGTG	1880	
QY	2234	GG	CTCAGCTT	CCTCTG	ACCGATG	AGGTAGTATAG	GAGACCTCAGT	GCCATCGATGAAA	2293	
Db	1881	GAG	TAGATGTG	AGCCTTGT	CATGGC	AGGGAGCTTGGG	AGACCTGTTTGGT	TGGAGGAGA	1940	
QY	2294	C	---CCCA	ATCATG	GACTCA	CAACATTTCCA-CTG	CAAC	CAGCCCTGTTTGC	CAAGCC	2349
Db	1941	CC	AGCCCGGTG	TAGACTCC	ACAGTATCAT	CCTGGC	AGCGAGTGTGCG	CTGCAACTC	2000	
QY	2350	AG	GAAACCCCGC	AGGGACC	AAATTTGGA---TGG	AGCACGTATAA	CACTAGAAAGCCTAAAA	2406		
Db	2001	CG	AAACACAG	CGTATCC	ACAGTGC	CAGAGACTTGG	AAACAGCTG	GAAGAAGAGAAATTAAG	2060	
QY	2407	CA	AGTCANT	GGATGTTTCC	CGGAC	TGAACCCCAT	CTACCCCT	TAAGCCCTCGGGTC	2466	
Db	2061	AC	ATGAA	AGATTCTCTG	CTCATTTATG	AAACCAATG	CCACTG	CCGAACACCCG-CTC	2119	
QY	2467	TC	CTCTCAT	AGGAAATGG	CACAGCTT	CAACACACCTG	CAGCTTGGT	TGGGCCCATG	2526	
Db	2120	CC	GACCAT	CAGTGC	AAATGGTTTC	GATGTA	TAGGTGCC	ACGAGTG-----TG	2170	
QY	2527	AC	GCTTCTCC	GGGCGAAG	AGACCTCTCTGGG	TGGACGTC	ATATGCTG	AAACCTGCTC	2586	
Db	2171	CT	ATCAACCC	CAGACATAG	AACTGTCC	ATGAATGAT	TGACCATG	ATGAACCAATTG	2230	
QY	2587	AA	CAGAA	GGCAGCAG	CGCCAGC	ACCATCAGCTT	CGGCTTAC	TGACGACCGCCGCTCG	2646	
Db	2231	AA	TGAGCG	CGGTGAC	AGACAC	CAGCAGC	ACCCTCAGCTC	AGCATAC	ACCGCCAGGCTCG	





/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6277469"  
/clone\_lib="NIH\_MGC\_102"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."  
BASE COUNT 202 a 253 c 258 g 157 t 1 others  
ORIGIN

Query Match 16.3%; Score 826; DB 14; Length 871;  
Best Local Similarity 97.9%; Pred. No. 8.4e-179;  
Matches 847; Conservative 0; Mismatches 16; Indels 2; Gaps 1;

QY 3789 TGAACAGCCCTGTAGGCCCCCGCAGTATGGAACTGTCTCAACAGCGCAGTGGCCCC 3848  
Db 1 TGAACAGCCCTGTAGGCCCCCGCAGTATGGAACTGTCTCAACAGCGCAGTGGCCCC 60  
QY 3849 TGGTGCACTCGACGGTGCCTGTGGTCCCGGATTCAAGCTCAAGCTGAAGAGCACCCC 3908  
Db 61 TGGTGCACTCGACGGTGCCTGTGGTCCCGGATTCAAGCTCAAGCTGAAGAGCACCCC 120  
QY 3909 CATGCAAGGAGCGGGGGCGCAGCTGAATTTCCGCTGCGGTAGCGCCCAATAGTACGC 3968  
Db 121 CATGCAAGGAGCGGGGGCGCAGCTGAATTTCCGCTGCGGTAGCGCCCAATAGTACGC 180  
QY 3969 TGGCAGCATGGTGAATGGCATCGAAGCAGACAGCAGTGGAGAGGGGTACCTGGCTCA 4028  
Db 181 TGGCAGCATGGTGAATGGCATCGAAGCAGACAGCAGTGGAGAGGGGTACCTGGCTCA 240  
QY 4029 CCAGCTCTCTCGGCGACAGCATGACGACCCCGGGGCGAGCCCGCCCGGTTCAGCAGATGCT 4088  
Db 241 CCAGCTCTCTCGGCGACAGCATGACGACCCCGGGGCGAGCCCGCCCGGTTCAGCAGATGCT 300  
QY 4089 TGGGCGAGATTAGTGTCTACATCAACATCAACATCAACATCAACATCAACATCAACATCA 4148  
Db 301 TGGGCGAGATTAGTGTCTACATCAACATCAACATCAACATCAACATCAACATCAACATCA 360  
QY 4149 AGGGGCTCAGCGATGGGCGAGCGCTCAAGCTTGGCAGTGTGTCAGGGGTACCGACCC 4208  
Db 361 AGGGGCTCAGCGATGGGCGAGCGCTCAAGCTTGGCAGTGTGTCAGGGGTACCGACCC 420  
QY 4209 ATGTGCGAGCTTTGGGGCGAGCGCGCGAGCGCTATGCGAGGGGACAGCCTTGCTCTGCA 4268  
Db 421 ATGTGCGAGCTTTGGGGCGAGCGCGCGAGCGCTATGCGAGGGGACAGCCTTGCTCTGCA 480  
QY 4269 GTGAGGACAGCTCAGTGACACAGTCAAGCTGCGAGGGTGAATGTTATCAAGATGGAGAT 4328  
Db 481 GTGAGGACAGCTCAGTGACACAGTCAAGCTGCGAGGGTGAATGTTATCAAGATGGAGAT 540  
QY 4329 GAAAGGCGAGCCCCATCCGCTGTCTTAATCTGAGAACTACTCTGCTGCTGCTGCTGCTG 4388  
Db 541 GAAAGGCGAGCCCCATCCGCTGTCTTAATCTGAGAACTACTCTGCTGCTGCTGCTGCTG 600  
QY 4389 CCAAAACCGTGGGCTTCAGTGACAGACACGAAAGCTGGTTCATCTTATTTACAGAGCC 4448  
Db 601 CCAAAACCGTGGGCTTCAGTGACAGACACGAAAGCTGGTTCATCTTATTTACAGAGCC 660  
QY 4449 CAGCTGCTGCTTACAGGGGACCGAGCCCAAACTCTGAGTTACTTTTCCCGAGGTGCTAA 4508  
Db 661 CAGCTGCTGCTTACAGGGGACCGAGCCCAAACTCTGAGTTACTTTTCCCGAGGTGCTAA 720  
QY 4509 TCAGGTGACAGCAGTGGAGCGCTCGACAGCCATGACCTGGAAGGGGTACAGATTGA 4568  
Db 721 TCAGGTGACAGCAGTGGAGCGCTCGACAGCCATGACCTGGAAGGGGTACAGATTGA 780

QY 4569 CTTGATGCCATCATAGACGATGGGACCACTCCAGCCTGA--TGTCGGGGGCCCTGAGC 4626  
Db 781 CTTGATGCCATCATAGACGATGGGACCACTCCAGCCTTAATGTCGGGGGCCCTGAGC 840  
QY 4627 CCAAGTATCATTCAGAACCTTTCC 4651  
Db 841 CCCAATATCATTCGGAACCCCTTCC 865  
RESULT 2  
BM552681 BM552681 998 bp mRNA linear EST 20-FEB-2002  
LOCUS AGENCOURT\_6573939 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5477180  
DEFINITION 5', mRNA sequence.  
ACCESSION BM552681  
VERSION BM552681.1 GI:18790743  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 998)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Cloning Strategy: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LICM1994 row: m column: 21  
High quality sequence stop: 668.  
FEATURES  
source  
1. 998  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5477180"  
/clone\_lib="NIH\_MGC\_98"  
/tissue\_type="astrocytoma grade IV, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 239 a 278 c 278 g 203 t  
ORIGIN  
Query Match 15.4%; Score 778; DB 13; Length 998;  
Best Local Similarity 92.8%; Pred. No. 9.3e-168;  
Matches 872; Conservative 0; Mismatches 60; Indels 8; Gaps 5;  
QY 3958 AATGAGTCAGTGGCAGCATGGTGAATGGCATGCAGAACCCAGTGGGACAGGG 4017  
Db 1 AATGAGTCAGTGGCAGCATGGTGAATGGCATGCAGAACCCAGTGGGACAGGG 60  
QY 4018 TACCTGCTCACCAGCTCTCGGGCAGCATGCAGACCCGGGGGAGCGCCCGCGT 4077  
Db 61 TACCTGCTCACCAGCTCTCGGGCAGCATGCAGACCCGGGGGAGCGCCCGCGT 120  
QY 4078 CAGCAGATGCTTGGGCGAGATTAGTGTACCTCACACATCAACATCTACCAAGGCCAGAG 4137  
Db 121 CAGCAGATGCTTGGGCGAGATTAGTGTACCTCACACATCAACATCTACCAAGGCCAGAG 180  
QY 4138 AGCTGCTGCCAGGGGCTCAGCGCATGGGCGAGCCGACGCGTCAAGCTTGGCAGTTGTCAGG 4197  
Db 4138 AGCTGCTGCCAGGGGCTCAGCGCATGGGCGAGCCGACGCGTCAAGCTTGGCAGTTGTCAGG 4197



Db 181 AGCTGCTGCTGAGGGGCTACGGCATGGGCGAGCGCTCAAGCTTGCGAGTTGTCAGG 240  
 QY 4198 GGCTACAGCCATGTCAGCTTTGGGGCAGCAGCGCGCAGCTATGCGGAGGACAGC 4257  
 Db 241 GGCTACAGCCATGTCAGCTTTGGGGCAGCAGCGCGCAGCTATGCGGAGGACAGC 300  
 QY 4258 CTTGCTCTGAGTCAGCAGCAGCTCAGTGACACAAGTCAGACCTGCGAGGCTGAATGATC 4317  
 Db 301 CTTGCTCTGAGTCAGCAGCAGCTCAGTGACACAAGTCAGACCTGCGAGGCTGAATGATC 360  
 QY 4318 AAGATGAGATGAAGAGCGAGCCCATCGCTGTGCTTAATCTGAGAAATCTCTGTT 4377  
 Db 361 AAGATGAGATGAAGAGCGAGCCCATCGCTGTGCTTAATCTGAGAAATCTCTGTT 420  
 QY 4378 CAGTTCTATCAACCAACCGTGGGCTTCAGTCAGCAGACAGCAAGAACTGTTCTCTCT 4437  
 Db 421 CAGTTCTATGACCAACCGTGGGCTTCAGTCAGCAGACAGCAAGAACTGTTCTCTCT 480  
 QY 4438 ATTTCAGACCCAGCTGCTGTACAGGGGACCGCCCAAAACCTCTGAGTTACTTTC 4497  
 Db 481 ATTTCAGACCCAGCTGCTGTACAGGGGACCGCCCAAAACCTCTGAGTTACTTTC 540  
 QY 4498 CCAGTGCTTAATCAGGTGACACAGCAGTGAGCAGCTGACAGCCATGACCTGGAAGG 4557  
 Db 541 CCAGTGCTTAATCAGGTGACACAGCAGTGAGCAGCTGACAGCCATGACCTGGAAGG 600  
 QY 4558 GTACAGATTGACTTCGATGCCATCATAGACGATGGGACCACTCCAGCCTGATGTCGGG 4617  
 Db 601 GTACAGATTGACTTCGATGCCATCATAGACGATGGGACCACTCCAGCCTGATGTCGGG 660  
 QY 4618 GCCTGAGCCCAAGTATCATTCAGAACCTTTCCTCATAGCTCCGCCCTCACCAGCCT 4677  
 Db 661 GCCTGAGCCCAAGTATCATTCAGAACCTTTCCTCATAGCTCCGCCCTCACCAGCCT 720  
 QY 4678 CGGGCGTCCCTCCATCCAGTCGCTGTCATGAGCAGCAGCAGCAACATGCTATC 4736  
 Db 721 CGGGCGTCCCTCCATCCAGTCGCTGTCATGAGCAGCAGCAGCAACATGCTATC 779  
 QY 4737 ACATGAGTCTTTGCTGAGCC--TCCCTAGCGGAAGAAAGCAATTC--TTGCGATTATG 4792  
 Db 780 ACATGAGTCTTTGCTGAGCCCTCCCTAGCGGGAAGAAAGCAATTCCTTGGCGGTTATG 839  
 QY 4793 CAATAGG--CTTTAGGAAAAAAGACTGCAACCAACCGAAATCAATAGAGTTGAAGAGA 4850  
 Db 840 CAATAGGCTTTAGGAAAAAAGACTGCAACCAACCGAAATTCCTAGGAGTTGGAA 899  
 QY 4851 TTAAGTGAATGTTGTTGCTGTTTGTGTTTGTGTTTGTGTTAT 4890  
 Db 900 AAGATTAAACCTGACTTTTGGTTTGGCTGTTTGTGTTTGTGTTT 939

RESULT 3  
 AA89686/c 890 bp mRNA linear EST 04-JAN-1999  
 LOCUS ak54c03.s1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:1409764  
 DEFINITION 3' similar to gb:M57609 GLI3 PROTEIN (HUMAN);, mRNA sequence.  
 ACCESSION AA89686  
 VERSION AA89686.1 GI:3016565  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 890)  
 NCBI-GenBank <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT: Robert Strausberg, Ph.D.  
 EMAIL: cgaplib-femail.nih.gov  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
 , Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.lnl.gov/bbrp/image/image.html  
 Insert Length: 995 Std Error: 0.00  
 Seq primer: -40ml3 fwd. ET from Amersham  
 High quality sequence stop: 498.

# FEATURES source

1..890  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1409764"  
 /clone\_lib="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - oligo(dT) primer [5',  
 TGTTCACCAATCTGAAGTGGGAGCGCCCAATTTTTTTTTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cots, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 224 a 196 c 219 g 248 t 3 others  
 ORIGIN  
 Query Match 14.1%; Score 714.6; DB 9; Length 890;  
 Best Local Similarity 97.0%; Pred. No. 3.3e-153;  
 Matches 811; Conservative 0; Mismatches 17; Indels 8; Gaps 8;

QY 4220 TTGGGGGAGCAGGCGCCAGGCTATGCGAGGAGCAGGCTGCTCTGAGTCAGCAGCAGC 4279  
 Db 890 TTGGGGGAGCAGGCGCCAGGCTATGCGGA--GGAGAGCCCTGCTCTGAGTCAGCAGCAGC 832  
 QY 4280 TCAGTGACACAAGTCAGACCTGCGAGGCTGAATGATCAAGATGAGATGAAGGGCAGC 4339  
 Db 831 TCA--TGACACAAGTCAGA--CTGAGGGTGAATGTTATCAAGATGAGATGAAGGGCAGC 774  
 QY 4340 CCCATCGCTGTGCTCTAATCTGCGAATTAATCTGCTGAGTTCTATGACCAAAACCGTGG 4399  
 Db 773 CCCATCGCTGTGCTCTAATCTGCGAATTAATCTGCTGAGTTCTATGACCAAAACCGTGG 714  
 QY 4400 GCTTCAGTCAGCAACAGCAGCAAGCTGTTCTATTCTTCTATTCAGACCCAGCCTGCCCTGC 4459  
 Db 713 GCTTCAGTCAGCAACAGCAGCAAGCTGTTCTATTCTATTTCAGACCCAGCCTGCCCTGC 654  
 QY 4460 TACAGGGACCAAGCGCCAAAACCTCTGAGTTACTTTCCCGAGGCTGCTAATCAGGTGACAA 4519  
 Db 653 TACA--GGGACCAAGCGCCAAAACCTCTGAGTTACTTTCCCGAGGCTGCTAATCAGGTGACAA 595  
 QY 4520 GCACAGTGACAGCCTGCGACGCAATGACCTGGAAGGGGTACAGATTGACTTTCGATGCCA 4579  
 Db 594 GCACAGTGACAGCCTGCGACGCAATGACCTGGAAGGGGTACAGATTGACTTTCGATGCCA 535  
 QY 4580 TCATAGCAGTGGGACCACTCCAGCCTGATGCTCGGGGCTGAGCCCAAGATATCATTC 4639  
 Db 534 TCATAGCAGTGGGACCACTCCAGCCTGATGTC--GGGGCCCTGAGCCCAAGATATCATTC 476  
 QY 4640 AGAAGCTTTCCCATAGCTCTCCCGCTCACACAGCCTCGGGGCTCCCTCCCATTTCCAG 4699  
 Db 475 AGAAGCTTTCCCATAGCTCTCCCGCTCACACAGCCTCGGGGCTCCCTCCCATTTCCAG 417  
 QY 4700 TCGCTGTCCATGAGCAGCAGCAACATGCGTATCGGGGATGAGTTCTTTGCTGACCTCC 4759  
 Db 416 -CGCTGTCCATGAGCAGCAGCAACATGCGTATCGGGGATGAGTTCTTTGCTGACCTCC 358  
 QY 4760 CTAGCGGAAGAAGCAAAATTCCTTCAGTTATGCAATAGGCTTTAGGAAAAAAGACATGC 4819  
 Db 357 CTAGCGGAAGAAGCAAAATTCCTTCAGTTATGCAATAGGCTTTAGG--AAAAAGACATGC 299  
 QY 4820 AACCAAGCAAAATCAATAGGAGTTGAAGAGATTAAGATGACTTCTTTTGGCTGTTT 4879

```

Db 298 AACCAACGAAATCAATAGAGTTGAAGAGATTAAACTGACTTTGTTGGCTGTTTTT 239
QY 4880 TAGTCTGTATGTTATTTAGCAATCTCATCTCACCTAAGTGTGTTCAATATAT 4939
Db 238 TAGTCTGTATGTTATTTAGCAATCTCATCTCACCTAAGTGTGTTCAATATAT 179
QY 4940 TCCTTTATGAAAAGGACTCTGAAAAACCCCTAAAGTATTCTAGGAGAAAGTGTCTTCC 4999
Db 178 TCCTTTATGAAAAGGACTCTGAAAAACCCCTAAAGTATTCTAGGAGAAAGTGTCTTCC 119
QY 5000 ATTTCAGTTTGAACAGTATGTTACACTCAACACACCTCTTTTAAAAA 5055
Db 118 ATTTCAGTTTGAACAGTATGTTACACTCAACACACCTCTTTTAAAAA 63

RESULT 4
AUI23517 763 bp mRNA linear EST 01-AUG-2002
LOCUS AUI23517 NT2RM2 Homo sapiens cDNA clone NT2RM2000468 5', mRNA
DEFINITION AUI23517 sequence.
ACCESSION AUI23517
VERSION AUI23517.1 GI:10948233
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 763)
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
1..763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM2000468"
/clone_lib="NT2RM2"
/cell_type="teratocarcinoma"
/cell_lines="NT2"
/notes="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 158 a 267 c 213 g 119 t 6 others
ORIGIN
1
Query Match 13.7%; Score 692.2; DB 9; Length 763;
Best Local Similarity 98.2%; Pred. No. 4.5e-148;
Matches 709; Conservative 0; Mismatches 11; Indels 2; Gaps 1;
QY 3199 TACACGCGCCCGAGGGGGCCAGTCCCGAAACTTCCACTGTCCCTGTCTCCCGAGC 3258
Db 1 TACACGCGCCCGAGGGGGCCAGTCCCGAAACTTCCACTGTCCCTGTCTCCCGAGC 60
QY 3259 ATCACCGAGAAGCTCACCTGGAGTCCCTGACCATGGAGCTGATGCCAACTGAACGAT 3318
Db 61 ATCACCGAGAAGCTCACCTGGAGTCCCTGACCATGGAGCTGATGCCAACTGAACGAT 120
QY 3319 GAGGATTTCTCCGCGAGACGCTGTGTGAGTATTAAATTTCCAGAACCAAGAGGTAC 3378

```

```

Db 121 GAGGATTTCTCCGCGAGACGCTGTGTGAGTATTAAATTTCCAGAACCAAGAGGTAC 180
QY 3379 GAGCAGACATTTCCCGAGCGGCTCCCGGAGCAGCAAGATGCCCCACGGCCGGTGAC 3438
Db 181 GAGCAGACATTTCCCGAGCGGCTCCCGGAGCAGCAAGATGCCCCACGGCCGGTGAC 240
QY 3439 TTTTACGCGCGCGGGCTGCCAGACGACGCTGCCAGCAGTTCCTATGCCCTCGAGCAG 3498
Db 241 TTTTACGCGCGCGGGCTGCCAGACGACGCTGCCAGCAGTTCCTATGCCCTCGAGCAG 300
QY 3499 CCCTGCCCCGAGGCGAGCAAAACCGACTGCCCATTCAGTGGAAACGAGTCAAGTCCGGGA 3558
Db 301 CCCTGCCCCGAGGCGAGCAAAACCGACTGCCCATTCAGTGGAAACGAGTCAAGTCCGGGA 360
QY 3559 AGCGCCGACCTGTCTCTCAAGCTCAAGTGTGGGCGGCGGCGCTGTGCCGAGACT 3618
Db 361 AGCGCCGACCTGTCTCTCAAGCTCAAGTGTGGGCGGCGGCGCTGTGCCGAGACT 420
QY 3619 CGCGCTTTGGGTTCTCAAGCGGCTGTCTCCACCGCAGAACCCCTTGAGGAGCGGG 3678
Db 421 CGCGCTTTGGGTTCTCAAGCGGCTGTCTCCACCGCAGAACCCCTTGAGGAGCGGG 480
QY 3679 CCTCTCGGGGCTATCAGACCTCTCGGGGAGAACCAACCCCTACGCTGGCCCGAGCAGC 3738
Db 481 CCTCTCGGGGCTATCAGACCTCTCGGGGAGAACCAACCCCTACGCTGGCCCGAGCAGC 540
QY 3739 TTGATGCTCCACACAGACCGCGGAGTGGCCACAGTGGAAACGCTTCCATGAACAGCCC 3798
Db 541 TTGATGCTCCACACAGACCGCGGAGTGGCCACAGTGGAAACGCTTCCATGAACAGCCC 600
QY 3799 TGTAAAGCCCGCGAGTATGGGAAGTCTCAACAGGAGCAGTGGCCCTGTGTGACTC 3858
Db 601 TGTAAAGCCCGCGAGTATGGGAAGTCTCAACAGGAGCAGTGGCCCTGTGTGACTC 660
QY 3859 GACGCTGCCTGTGTGCGGGGATTCATCAAGAGTGAAGAGCAGCCCCCATGCAAG 3916
Db 661 GACGCTGCCTGTGTGCGGGGATTCATCAAGAGTGAAGAGCAGCCCCCATGCAAG 720
QY 3917 GG 3918
Db 721 GG 722

RESULT 5
BI830216 702 bp mRNA linear EST 04-OCT-2001
LOCUS 603072918F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164757 5',
DEFINITION mRNA sequence.
ACCESSION BI830216
VERSION BI830216.1 GI:15941766
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 702)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11409 row: d column: 06
High quality sequence stop: 695.
Location/Qualifiers
1..702
FEATURES
source

```





/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4415095"  
 /clone\_lib="NIH\_MGC\_89"  
 /tissue\_type="hypernephroma, cell line"  
 /lab\_host="PH10B (phage-resistant)"  
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.3 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC library."

BASE COUNT 250 a 246 c 233 g 205 t  
 ORIGIN

Query Match 13.1%; Score 660.6; DB 12; Length 934;  
 Best Local Similarity 97.3%; Pred. No. 8.6e-141;  
 Matches 715; Conservative 0; Mismatches 14; Indels 6; Gaps 4;

QY 4200 CTACAGCCATGTGCCAGCTTTTGGGGGACGAGCGCCAGGCTATGCCGAGGGACAGCCT 4259  
 Db 1 CTACAGCCATGTGCCAGCTTTTGGGGGACGAGCGCCAGGCTATGCCGAGGGACAGCCT 60

QY 4260 TGCTCTGAGTCAGGAGCTAGTGCAGACAGTGCAGAGCTGCAGGCTGAATGGTATCAA 4319  
 Db 61 TGCTCTGAGTCAGGAGCTAGTGCAGACAGTGCAGAGCTGCAGGCTGAATGGTATCAA 120

QY 4320 GATGGAGATGAAGGGGACGCGCCATCGCTGTCTCTAATCTGCAGAAATCTCTGGTCA 4379  
 Db 121 GATGGAGATGAAGGGGACGCGCCATCGCTGTCTCTAATCTGCAGAAATCTCTGGTCA 180

QY 4380 GTTCTATGACAAACCGTGGGCTTCAGTCAGCAAGACAGAAAGCTGGTTCTATCTCTAT 4439  
 Db 181 GTTCTATGACAAACCGTGGGCTTCAGTCAGCAAGACAGAAAGCTGGTTCTATCTCTAT 240

QY 4440 TTCAGAGCCAGCTGCTGCTCAGGGGACGAGCGCCAGCAAACTCTGAGTTCTATTTCCCC 4499  
 Db 241 TTCAGAGCCAGCTGCTGCTCAGGGGACGAGCGCCAGCAAACTCTGAGTTCTATTTCCCC 300

QY 4500 AGTGCTAATCAGGTGACAGACAGTGGACAGCCTTCGACAGCATGACCTGGAAGGGT 4559  
 Db 301 AGTGCTAATCAGGTGACAGACAGTGGACAGCCTTCGACAGCATGACCTGGAAGGGT 360

QY 4560 ACAGATTGACTTCGATGCCATCATACAGATGGGGACCACTCCAGCCTGATGTCGGGGGC 4619  
 Db 361 ACAGATTGACTTCGATGCCATCATACAGATGGGGACCACTCCAGCCTGATGTCGGGGGC 420

QY 4620 CCTGAGCCCAAGTATCATTCAGAACCTTCCCATAGCTCCTCCCGCCTCACCAGCCTCG 4679  
 Db 421 CCTGAGCCCAAGTATCATTCAGAACCTTCCCATAGCTCCTCCCGCCTCACCAGCCTCG 480

QY 4680 GCGGTCCCTCCCATCCAGTCGCTGCTCATGAGCACCACCAATGCGTATCGGGGACA 4739  
 Db 481 GCGGTCCCTCCCATCCAGTCGCTGCTCATGAGCACCACCAATGCGTATCGGGGACA 539

QY 4740 TGAGTTCTTTGCTGACCTCCCTAGCGGAAGAAAGCAAAATCTTTCAGTATGCAATAGG 4799  
 Db 540 TGAGTTCTTTGCTGACCTCCCTAGCGGAAGAAAGCAAAATCTTTCAGTATGCAATAGG 599

QY 4800 CTTTAGAAAAAAGACTGCAACCAACGGAATCAATAGAGTTGAAGAGATTAACCTGA 4859  
 Db 600 CTTTAGAAAAAAGACTGCAACCAACGGAATCAATAGAGTTGAAGAGATTAACCTGA 656

QY 4860 CTTTGTGTTGGCTGTTTTTTAGTTCTGTATGTTATTTTACCAATCTCCTACCTAAT 4919  
 Db 657 CTTTGTGTTGGCTGTTTTTTAGTTCTGTATGTTATTTTACCAATCTCCTACCTAAT 714

QY 4920 GAGATGTTTCAAT 4934  
 Db 715 GAGATGTTTCAAT 729

RESULT 9  
 AUI20131

LOCUS AUI20131 713 bp mRNA linear EST 01-AUG-2002  
 DEFINITION AUI20131 HEMBB1 Homo sapiens cDNA clone HEMBB1000032 5', mRNA  
 sequence.  
 ACCESSION AUI20131  
 VERSION AUI20131  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 713)  
 Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,  
 Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and  
 Isogai, T.  
 HRI human cDNA project  
 Unpublished (2000)  
 Contact: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute.  
 Location/Qualifiers  
 1. 713  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HEMBB1000032"  
 /clone\_lib="HEMBB1"  
 /tissue\_type="whole embryo, mainly body"  
 /dev\_stage="embryo, 10 weeks"  
 /note="Vector: pME18SFL3"

BASE COUNT 142 a 235 c 222 g 111 t 3 others  
 ORIGIN

Query Match 12.7%; Score 642.8; DB 9; Length 713;  
 Best Local Similarity 99.7%; Pred. No. 9.9e-137;  
 Matches 644; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3365 ACCAAGCAGGTTACGAGCAGCAGCTTCCCGAGCGCCCTCCCGAGCAGCAGCAAGTGCCTC 3424  
 Db 1 ACCAAGCAGGTTACGAGCAGCAGCTTCCCGAGCGCCCTCCCGAGCAGCAGCAAGTGCCTC 60

QY 3425 AGGGCCCGGTGACTTTTACGCGCCCGGCTGCCAGACAGCAGCAGCTGCCAGCAGTTCC 3484  
 Db 61 AGGGCCCGGTGACTTTTACGCGCCCGGCTGCCAGACAGCAGCAGCTGCCAGCAGTTCC 120

QY 3485 ATGCCCTCAGAGCGCTTCCCGGAGGCGAGCAAAACCGACCTGCCCATTCAGTGAACG 3544  
 Db 121 ATGCCCTCAGAGCGCTTCCCGGAGGCGAGCAAAACCGACCTGCCCATTCAGTGAACG 180

QY 3545 AGTCAGCTCCGAGAGCGCGACCTGCTCTCTCAAGCTCAAGTGTGGGCGCGGGCCG 3604  
 Db 181 AGTCAGCTCCGAGAGCGCGACCTGCTCTCTCTCAAGCTCAAGTGTGGGCGCGGGCCG 240

QY 3605 CTGTGCCCGCAGACTCGCGCTTTTGGGTTCTGCAACCGCATGCTGCTCCACCGCAGAAC 3664  
 Db 241 CTGTGCCCGCAGACTCGCGCTTTTGGGTTCTGCAACCGCATGCTGCTCCACCGCAGAAC 300

QY 3665 CTTGAGGAGCGGGCTGCTGGGGCTATCAGACCTTCGGGGAGAACCAACCCCTACG 3724  
 Db 301 CTTGAGGAGCGGGCTGCTGGGGCTATCAGACCTTCGGGGAGAACCAACCCCTACG 360

QY 3725 GTGGCCCAAGCAGCTTGATGCTCCACACAGCCCCGGAGTGGCACCAGTGAACACGCT 3784  
 Db 361 GTGGCCCAAGCAGCTTGATGCTCCACACAGCCCCGGAGTGGCACCAGTGAACACGCT 420

QY 3785 TCCATGAACAGCCCTGTAAAGCCCGCAGTATGGGAACCTGCTCAACAGGCGCAGTGG 3844  
 Db 3844



/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3367412"  
 /clone\_lib="NCI CGAP\_Lu24"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP\_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 141920-141791 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 179 a 142 c 173 g 184 t  
 ORIGIN

Query Match 12.5%; Score 629.4; DB 12; Length 678;  
 Best Local Similarity 97.3%; Pred. No. 1.2e-133;  
 Matches 661; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

QY 4372 TCTGTCAGTTCCTATGACCAACCGT-GGGCTTACGACGACAGACAGAAAGCTGCTTC 4430  
 |||||||  
 Db 678 TCTGTCAGTTCCTATGACCAACCGTGGCTTACGACGACAGACAGAAAGCTGCTTC 619  
 |||||||  
 QY 4431 ATTCTATTACAGGCGACGCTGCTGCTACAGGGGACGCGCCAAACCTGAGTT 4490  
 |||||||  
 Db 618 ATTCTATTACAGGCGACGCTGCTGCTACAGGGGACGCGCCAAACCTGAGTT 559  
 |||||||  
 QY 4491 ACTTCCCGAGGCTCTATACAGTGCAGACAGACAGTGGAGCCTCGACAGCCATGACCT 4550  
 |||||||  
 Db 558 ACTTCCCGAGGCTCTATACAGTGCAGACAGACAGTGGAGCCTCGACAGCCATGACCT 499  
 |||||||  
 QY 4551 GGAAGGGGTACAGATTGACCTTCGATCAGATAGAGATGGGACCACTCCACCTGAT 4610  
 |||||||  
 Db 498 GGAAGGGGTACAGATTGACCTTCGATCAGATAGAGATGGGACCACTCCACCTGAT 439  
 |||||||  
 QY 4611 GTGGGGGGCCCTGAGCCCAAGTATCATTGACAGCTTCCCATAGCTCCCGCCCTCAC 4670  
 |||||||  
 Db 438 GTGGGGGGCCCTGAGCCCAAGTATCATTGACAGCTTCCCATAGCTCCCGCCCTCAC 379  
 |||||||  
 QY 4671 CAGCGCTGGGGCTCCCTCCATCCAGTGCCTGCTCCATGAGCAGCAGCAACATGGCTA 4730  
 |||||||  
 Db 378 CAGCGCTGGGGCTCCCTCCATCCAGTGCCTGCTCCATGAGCAGCAGCAACATGGCTA 320  
 |||||||  
 QY 4731 TCGGGGACATGATCTTCTGCTGACCTCCCTAGCGGAAGAAAGCAATTCCTGCACTTA 4790  
 |||||||  
 Db 319 TCGGGGACATGATCTTCTGCTGACCTCCCTAGCGGAAGAAAGCAATTCCTGCACTTA 260  
 |||||||  
 QY 4791 TGCATATAGCTTTAGGAAAAAAGACACGCAACCAAGCAATTCATAGGATGAAGAGA 4850  
 |||||||  
 Db 259 TGCATATAGCTTTAGGAAAAAAGACACGCAACCAAGCAATTCATAGGATGAAGAGA 200  
 |||||||  
 QY 4851 TTTAACTGACTTTTGGCTGTTTTTTTGTAGTCTGATGATTTTGAAGATCTCATCT 4910  
 |||||||  
 Db 199 TTTAACTGACTTTTGGCTGTTTTTTTGTAGTCTGATGATTTTGAAGATCTCATCT 140  
 |||||||  
 QY 4911 CACCTACTGAGATGTGTTCAATTAATTCCTTTTATGAAAGAGACTGGAAGAACCC 4970  
 |||||||  
 Db 139 CACCTACTGAGATGTGTTCAATTAATTCCTTTTATGAAAGAGACTGGAAGAACCC 80  
 |||||||  
 QY 4971 TAAATATTTAGGAGAAAGTGTCTTCATTTGAGTTTAAACAGATGTTACACTC 5030  
 |||||||  
 Db 79 TAAATATTTAGGAGAAAGTGTCTTCATTTGAGTTTAAACAGATGTTACACTC 20  
 |||||||  
 QY 5031 AAACCACTCTTTTAA 5049  
 |||||||  
 Db 19 AAACCACTCTTTTAA 1

RESULT 12  
 B0447085/c

LOCUS B0447085 694 bp mRNA linear EST 29-MAY-2002  
 DEFINITION UI-H-EU1-bad-g-19-0-UI-s1 NCI-CGAP-Cl1 Homo sapiens CDNA clone  
 UI-H-EU1-bad-g-19-0-UI 3', mRNA sequence.  
 ACCESSION B0447085  
 VERSION B0447085.1 GI:21250197  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 694)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Jose Mercuende  
 CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, Bento-soares@uiowa.edu  
 The following repetitive elements were found in this CDNA  
 sequence: 61-83, >AT-rich#Low-complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..694  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-EU1-bad-g-19-0-UI"  
 /clone\_lib="NCI CGAP Cl1"  
 /tissue\_type="Osteoarthritic Cartilage"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Organ: Knee; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Cl1 is a normalized CDNA library containing the following tissue(s): Osteoarthritic Cartilage The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCGT.  
 TAG\_LIB=UI-H-EU1  
 TAG\_TISSUE=osteoarthritis cartilage  
 TAG\_SEQ=TGATCAGCGT"

BASE COUNT 174 a 138 c 171 g 210 t 1 others  
 ORIGIN

Query Match 11.8%; Score 598.4; DB 14; Length 694;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-126;  
 Matches 610; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 4443 AGACGGCAGTGGCTGCTACAGGGGACGACGCGCAAAACTGAGTACTTCCCAAG 4502  
 |||||||  
 Db 681 AGGCGCAGTGGCTGCTACAGGGGACGACGCGCAAAACTGAGTACTTCCCAAG 622  
 |||||||  
 QY 4503 TGCTAATCAGGTGACAGCAAGCACTGACAGCTCGACAGCCATGACCTGGAGGGGTACA 4562  
 |||||||  
 Db 621 TGCTAATCAGGTGACAGCAAGCACTGACAGCTCGACAGCCATGACCTGGAGGGGTACA 562  
 |||||||  
 QY 4563 GATTGACTTGAGTGCATCATACATGAGGAGCACTCAGGCTGATGGGGGGCCCT 4622  
 |||||||  
 Db 561 GATTGACTTGAGTGCATCATACATGAGGAGCACTCAGGCTGATGGGGGGCCCT 502  
 |||||||  
 QY 4623 GAGCCCAAGTATCAGAACCTTCCCATAGCTCTCCCGCTCAGACAGCGCTCGGGC 4682



Db 501 GAGCCCAAGATCATTCAGACCTTCCCATAGCTCTCCGCTCAGCAGCGCTGGGC 442

Qy 4683 GTCCCTCCCATTCGCCAGTGTCTCATGACACACCAACATGGCTATGGGACATGA 4742

Db 441 GTCCCTCCCATTCGCCAGTGTCTCATGACACACCAACATGGCTATGGGACATGA 383

Qy 4743 GTTCCTTCGACCTCCCTAGGGAAGAAAGCAATTCCTTCAGTATGCAATGAGCTT 4802

Db 382 GTTCCTTCGACCTCCCTAGGGAAGAAAGCAATTCCTTCAGTATGCAATGAGCTT 323

Qy 4803 TAGGAAAAAAGAGCTCAACCAAGCAATTCATAGAGCTTGAAGATTAAGTACTT 4862

Db 322 TAGGAAAAAAGAGCTCAACCAAGCAATTCATAGAGCTTGAAGATTAAGTACTT 263

Qy 4863 TGTTCCTTCGACCTCCCTAGGGAAGAAAGCAATTCCTTCAGTATGCAATGAGCTT 4922

Db 262 TGTTCCTTCGACCTCCCTAGGGAAGAAAGCAATTCCTTCAGTATGCAATGAGCTT 203

Qy 4923 ATGTGTTCAATATATATTCCTTTATGAAAAAGAGCTCGAAAAACCTTAAGTATCTA 4982

Db 202 ATGTGTTCAATATATATTCCTTTATGAAAAAGAGCTCGAAAAACCTTAAGTATCTA 143

Qy 4983 GGGAGAACTGTCTTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 5042

Db 142 GGGAGAACTGTCTTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 83

Qy 5043 TTTTAAAAA 5055

Db 82 TTTTAAAAA 70

RESULT 13  
BE221252/c 576 bp mRNA linear EST 03-JUL-2000  
LOCUS hu22f05.x1 NCI-CGAP\_Mel15 Homo sapiens CDNA clone IMAGE:3170817 3'  
DEFINITION similar to gb:M57609 GLI3 PROTEIN (HUMAN);, mRNA sequence.  
ACCESSION BE221252  
VERSION BE221252.1 GI:8908570  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 576)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov

JOURNAL COMMENT  
REFERENCE  
AUTHORS  
TITLE  
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Christina Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov  
Seq primer: -400P from Glibco  
High quality sequence stop: 404.  
Location/Qualifiers  
1. 576  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3170817"  
/clone\_lib="NCI-CGAP\_Mel15"  
/tissue\_type="malignant melanoma, metastatic to lymph node"  
/lab\_host="DH10B"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site\_1: SalI; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies."

BASE COUNT 156 a 118 c 144 g 156 t 2 others  
ORIGIN

Query Match 11.1%; Score 563; DB 10; Length 576;  
Best Local Similarity 99.5%; Pred. No. 2e-118;  
Matches 574; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 4471 AGGCGCAAAATCTGATCTTCTCCAGTCTCTATATGATGACACAGACAGTGGAC 4530

Db 576 AGGCGCAAAATCTGATCTTCTCCAGTCTCTATATGATGACACAGACAGTGGAC 517

Qy 4531 AGCCTCGACACCATGACCTGGAAGGGGTACAGATTACTTCATCATATAGAGAT 4590

Db 516 AGCCTCGACACCATGACCTGGAAGGGGTACAGATTACTTCATCATATAGAGAT 457

Qy 4591 GGGGACCACTCCAGCTGATGTCGGGGCCCTGAGCCCAAGTATCATTCAGAACCTTCC 4650

Db 456 GGGGACCACTCCAGCTGATGTCGGGGCCCTGAGCCCAAGTATCATTCAGAACCTTCC 397

Qy 4651 CATAGCTCTCCGCTCAGCAGCGCTCGGGGCTCCCTCCATTCAGTCGCTTCAT 4710

Db 396 CATAGCTCTCCGCTCAGCAGCGCTCGGGGCTCCCTCCATTCAGTCGCTTCAT 338

Qy 4711 GAGCACCAACCAATGGCTATCGGGACATGATGTTCTTGTGACCTCCCTAGGGAAGA 4770

Db 337 GAGCACCAACCAATGGCTATCGGGACATGATGTTCTTGTGACCTCCCTAGGGAAGA 278

Qy 4771 AAGCAATTCCTTCGAGTTATGCAATAGCTTTAGAAAAAAGACGTGCAACGAGAA 4830

Db 277 AAGCAATTCCTTCGAGTTATGCAATAGCTTTAGAAAAAAGACGTGCAACGAGAA 218

Qy 4831 ATCAATAGAGTTGAAGAGATTAAGATGATCTTGTGCTTTTGTGTTTATGTTCTAT 4890

Db 217 ATCAATAGAGTTGAAGAGATTAAGATGATCTTGTGCTTTTGTGTTTATGTTCTAT 158

Qy 4891 GTATTTAGCAATGCTCATCTGACCTTAAGAGTGTTCATATATTCCTTTATGG 4950

Db 157 GTATTTAGCAATGCTCATCTGACCTTAAGAGTGTTCATATATTCCTTTATGG 98

Qy 4951 AAAAGACTCTGAAAAAACCTTAAGATTTAGGGAGAAACTGTCTTCATTCAGTTTT 5010

Db 97 AAAAGACTCTGAAAAAACCTTAAGATTTAGGGAGAAACTGTCTTCATTCAGTTTT 38

Qy 5011 GAATCAGTATGTTTACACTCAACCAACCCCTTTTGA 5047

Db 37 GAATCAGTATGTTTACACTCAACCAACCCCTTTTGA 1

RESULT 14  
A1796831 650 bp mRNA linear EST 18-DEC-1999  
LOCUS A1796831/c  
DEFINITION we22h03.x1 NCI-CGAP\_Lu24 Homo sapiens CDNA clone IMAGE:2341877 3'  
similar to gb:M57609 GLI3 PROTEIN (HUMAN);, mRNA sequence.  
ACCESSION A1796831  
VERSION A1796831.1 GI:5362294  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 650)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov

JOURNAL COMMENT  
REFERENCE  
AUTHORS  
TITLE  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be



```
Db 334 GCACCAACAATGGCTATCGGGACATGATGTTCTTTCCTGACCTCCCTAGCGAAGAAA 275
QY 4773 GCAAAATTCCTGCGATTATGCAATAGGCTTTAGGAAAAAAGACTGCACCAACGGAAT 4832
Db 274 GCAAAATTCCTGCGATTATGCAATAGGCTTTAGGAAAAAAGACTGCACCAACGGAAT 215
QY 4833 CATAGGAGTTGAAGAGATTAACTGACTTGTGCTGCTG-TTTTWTAGTTCGTATG 4891
Db 214 CATAGGAGTTGAAGAGATTAACTGACTTGTGCTGCTGTTTGTAGTCTGTATG 155
QY 4892 TATTTAGCAATCTCATCTCACCCTAAGATGCTTCAATTAATTCCTTTATGGA 4951
Db 154 TATTTAGCAATCTCATCTCACCCTAAGATGCTTCAATTAATTCCTTTATGGA 95
QY 4952 AAAGACTCTGAAAAACCTTAAGATTTCTAGGAGAAACTGCTTCATTCAGTTTG 5011
Db 94 AAAGACTCTGAAAAACCTTAAGATTTCTAGGAGAAACTGCTTCATTCAGTTTG 35
QY 5012 AATCAGTATTTGTACACTCAACCAACCTCTTTT 5045
Db 34 AATCAGTATTTGTACACTCAACCAACCTCTTTT 1
```

Search completed: December 11, 2002, 00:29:40  
Job time : 4311 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2002, 02:29:06 ; Search time 110 Seconds  
(without alignments)  
14093.191 Million cell updates/sec

Title: US-09-910-185-3  
Perfect score: 5055  
Sequence: 1 cgatactacgtggcattt.....acctcttttaaaaaaa 5055

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues  
Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : Issued\_Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29.6	0.6	78	4	US-09-043-303-11
2	28	0.6	69	4	US-09-043-303-13
3	28	0.6	78	4	US-09-043-303-12
4	27.4	0.5	99	1	US-08-240-049B-10
5	27.4	0.5	99	4	US-08-542-634-12
6	27.4	0.5	99	4	US-08-477-292-12
7	27.4	0.5	99	5	PCT-US95-13703-12
8	27	0.5	96	4	US-09-242-095-26
9	27	0.5	99	4	US-09-043-303-10
10	26.4	0.5	75	4	US-08-556-978B-67
11	25.8	0.5	93	4	US-08-556-978B-68
12	25.8	0.5	93	4	US-08-556-978B-68
13	25.6	0.5	84	1	US-08-405-702A-7
14	25.4	0.5	90	4	US-08-556-978B-31
15	25	0.5	90	4	US-08-556-978B-30
16	24.8	0.5	84	1	US-08-039-778B-7
17	24.6	0.5	66	4	US-08-556-978B-41
18	24.4	0.5	72	1	US-08-053-131-132
19	24.4	0.5	72	1	US-08-645-641-132
20	24.4	0.5	72	1	US-07-853-408B-132
21	24.4	0.5	72	1	US-08-096-762-132
22	24.4	0.5	72	2	US-08-308-865-132
23	24.4	0.5	72	5	PCT-US92-10983-132
24	24.2	0.5	57	4	US-09-043-303-14
25	24.2	0.5	69	1	US-08-039-778B-8
26	24.2	0.5	99	1	US-08-240-049B-9
27	24.2	0.5	99	4	US-08-542-634-11

28	24.2	0.5	99	4	US-08-477-292-11	Sequence 11, Appl
29	24.2	0.5	99	5	PCT-US95-13703-11	Sequence 11, Appl
30	24	0.5	69	1	US-08-053-131-129	Sequence 129, App
31	24	0.5	69	1	US-08-645-641-129	Sequence 129, App
32	24	0.5	69	1	US-07-853-408B-129	Sequence 129, App
33	24	0.5	69	1	US-08-096-762-129	Sequence 129, App
34	24	0.5	69	1	US-08-308-865-129	Sequence 129, App
35	24	0.5	69	5	PCT-US92-10983-129	Sequence 129, App
36	24	0.5	72	4	US-09-042-353-249	Sequence 249, App
37	24	0.5	72	4	US-08-758-417A-97	Sequence 97, App
38	24	0.5	84	1	US-08-405-702A-5	Sequence 5, Appl
39	24	0.5	84	1	US-08-405-702A-9	Sequence 9, Appl
40	24	0.5	87	2	US-08-476-176B-22	Sequence 22, Appl
41	24	0.5	87	3	US-08-127-721A-22	Sequence 22, Appl
42	24	0.5	87	3	US-08-485-246A-22	Sequence 22, Appl
43	24	0.5	100	1	US-08-145-705A-27	Sequence 27, Appl
44	24	0.5	100	1	US-08-145-705A-30	Sequence 30, Appl
45	24	0.5	45	4	US-08-339-214-85	Sequence 85, Appl
46	23.8	0.5	61	3	US-09-023-228B-35	Sequence 35, Appl
47	23.8	0.5	61	4	US-09-163-025B-35	Sequence 35, Appl
48	23.8	0.5	90	4	US-09-364-707A-16	Sequence 16, Appl
49	23.8	0.5	96	3	US-07-876-288-21	Sequence 21, Appl
50	23.6	0.5	79	4	US-09-143-211-9	Sequence 9, Appl
51	23.6	0.5	80	4	US-09-284-627-15	Sequence 15, Appl
52	23.4	0.5	71	2	US-08-712-948-4	Sequence 4, Appl
53	23.4	0.5	75	1	US-08-053-131-150	Sequence 150, App
54	23.4	0.5	75	1	US-08-645-641-150	Sequence 150, App
55	23.4	0.5	75	1	US-07-853-408B-150	Sequence 150, App
56	23.4	0.5	75	1	US-08-096-762-150	Sequence 150, App
57	23.4	0.5	75	2	US-08-308-865-150	Sequence 150, App
58	23.4	0.5	75	4	US-09-042-353-267	Sequence 267, App
59	23.4	0.5	75	4	US-08-758-417A-115	Sequence 115, App
60	23.4	0.5	75	5	PCT-US92-10983-150	Sequence 150, App
61	23.4	0.5	76	4	US-09-042-353-260	Sequence 260, App
62	23.4	0.5	76	4	US-08-758-417A-108	Sequence 108, App
63	23.4	0.5	77	1	US-08-053-131-143	Sequence 143, App
64	23.4	0.5	77	1	US-08-645-641-143	Sequence 143, App
65	23.4	0.5	77	1	US-07-853-408B-143	Sequence 143, App
66	23.4	0.5	77	1	US-08-096-762-143	Sequence 143, App
67	23.4	0.5	77	2	US-08-308-865-143	Sequence 143, App
68	23.4	0.5	77	2	PCT-US92-10983-143	Sequence 143, App
69	23.4	0.5	80	1	US-08-053-131-136	Sequence 136, App
70	23.4	0.5	80	1	US-08-645-641-136	Sequence 136, App
71	23.4	0.5	80	1	US-07-853-408B-136	Sequence 136, App
72	23.4	0.5	80	1	US-08-096-762-136	Sequence 136, App
73	23.4	0.5	80	2	US-08-308-865-136	Sequence 136, App
74	23.4	0.5	80	4	US-09-042-353-253	Sequence 253, App
75	23.4	0.5	80	4	US-08-758-417A-101	Sequence 101, App
76	23.4	0.5	80	5	PCT-US92-10983-136	Sequence 136, App
77	23.4	0.5	81	1	US-08-053-131-130	Sequence 130, App
78	23.4	0.5	81	1	US-08-645-641-130	Sequence 130, App
79	23.4	0.5	81	1	US-07-853-408B-130	Sequence 130, App
80	23.4	0.5	81	1	US-08-096-762-130	Sequence 130, App
81	23.4	0.5	81	2	US-08-308-865-130	Sequence 130, App
82	23.4	0.5	81	4	US-09-042-353-247	Sequence 247, App
83	23.4	0.5	81	4	US-08-758-417A-95	Sequence 95, Appl
84	23.4	0.5	81	5	PCT-US92-10983-130	Sequence 130, App
85	23.4	0.5	84	1	US-08-053-131-140	Sequence 140, App
86	23.4	0.5	84	1	US-08-053-131-145	Sequence 145, App
87	23.4	0.5	84	1	US-08-645-641-140	Sequence 140, App
88	23.4	0.5	84	1	US-08-645-641-145	Sequence 145, App
89	23.4	0.5	84	1	US-07-853-408B-145	Sequence 145, App
90	23.4	0.5	84	1	US-07-853-408B-145	Sequence 145, App
91	23.4	0.5	84	1	US-08-096-762-140	Sequence 140, App
92	23.4	0.5	84	1	US-08-096-762-145	Sequence 145, App
93	23.4	0.5	84	2	US-08-308-865-140	Sequence 140, App
94	23.4	0.5	84	2	US-08-308-865-145	Sequence 145, App
95	23.4	0.5	84	2	US-08-477-257A-141	Sequence 141, App
96	23.4	0.5	84	3	US-08-481-710-141	Sequence 141, App
97	23.4	0.5	84	4	US-09-042-353-257	Sequence 257, App
98	23.4	0.5	84	4	US-09-042-353-262	Sequence 262, App
99	23.4	0.5	84	4	US-08-758-417A-105	Sequence 105, App
100	23.4	0.5	84	4	US-08-758-417A-110	Sequence 110, App

101	23.4	0.5	84	5	PCT-US92-10983-140	Sequence 140, App	c 174	23	0.5	97	3	US-08-825-852-45	Sequence 45, Appl
102	23.4	0.5	84	5	PCT-US92-10983-145	Sequence 145, App	c 175	23	0.5	97	4	US-09-052-888-39	Sequence 39, Appl
103	23.4	0.5	84	5	PCT-US96-09537-141	Sequence 141, App	c 176	23	0.5	97	4	US-09-052-888-40	Sequence 40, Appl
104	23.4	0.5	85	1	US-08-053-131-148	Sequence 148, App	c 177	23	0.5	97	4	US-09-052-888-43	Sequence 43, Appl
105	23.4	0.5	85	1	US-08-645-641-148	Sequence 148, App	c 178	23	0.5	97	4	US-09-052-888-46	Sequence 46, Appl
106	23.4	0.5	85	1	US-07-853-408B-148	Sequence 148, App	c 179	23	0.5	99	1	US-07-918-953-3	Sequence 3, Appl
107	23.4	0.5	85	1	US-08-096-762-148	Sequence 148, App	c 180	23	0.5	99	1	US-07-918-953-3	Sequence 3, Appl
108	23.4	0.5	85	2	US-08-308-865-148	Sequence 148, App	c 181	23	0.5	99	1	US-08-081-661-3	Sequence 4, Appl
109	23.4	0.5	85	4	US-09-042-353-265	Sequence 265, App	c 182	23	0.5	99	1	US-08-081-661-3	Sequence 4, Appl
110	23.4	0.5	85	4	US-08-758-417A-113	Sequence 113, App	c 183	22.8	0.5	50	1	US-08-171-389-256	Sequence 256, App
111	23.4	0.5	85	5	PCT-US92-10983-148	Sequence 148, App	c 184	22.8	0.5	50	1	US-08-171-389-256	Sequence 256, App
112	23.4	0.5	87	1	US-08-053-131-149	Sequence 149, App	c 185	22.8	0.5	50	1	US-08-123-936-256	Sequence 256, App
113	23.4	0.5	87	1	US-08-645-641-149	Sequence 149, App	c 186	22.8	0.5	50	1	US-08-123-936-256	Sequence 256, App
114	23.4	0.5	87	1	US-07-853-408B-149	Sequence 149, App	c 187	22.8	0.5	50	2	US-08-475-228A-256	Sequence 256, App
115	23.4	0.5	87	2	US-08-096-762-149	Sequence 149, App	c 188	22.8	0.5	50	2	US-08-475-228A-256	Sequence 256, App
116	23.4	0.5	87	2	US-08-308-865-149	Sequence 149, App	c 189	22.8	0.5	50	3	US-08-482-080A-256	Sequence 256, App
117	23.4	0.5	87	4	US-09-042-353-266	Sequence 266, App	c 190	22.8	0.5	50	3	US-08-482-080A-256	Sequence 256, App
118	23.4	0.5	87	5	US-08-758-417A-114	Sequence 114, App	c 191	22.8	0.5	50	3	US-08-482-080A-256	Sequence 256, App
119	23.4	0.5	87	5	PCT-US92-10983-149	Sequence 149, App	c 192	22.8	0.5	50	4	US-08-482-080A-256	Sequence 256, App
120	23.4	0.5	90	1	US-08-053-131-131	Sequence 131, App	c 193	22.8	0.5	50	4	US-09-354-947-256	Sequence 256, App
121	23.4	0.5	90	1	US-08-645-641-131	Sequence 131, App	c 194	22.8	0.5	50	5	PCT-US93-12388-256	Sequence 256, App
122	23.4	0.5	90	1	US-07-853-408B-131	Sequence 131, App	c 195	22.8	0.5	50	5	PCT-US93-12388-256	Sequence 256, App
123	23.4	0.5	90	1	US-08-096-762-131	Sequence 131, App	c 196	22.8	0.5	62	2	US-07-812-420-25	Sequence 25, Appl
124	23.4	0.5	90	2	US-08-308-865-131	Sequence 131, App	c 197	22.8	0.5	62	2	US-07-812-420-25	Sequence 25, Appl
125	23.4	0.5	90	2	US-08-645-641-131	Sequence 131, App	c 198	22.8	0.5	65	4	US-09-415-784-32	Sequence 32, Appl
126	23.4	0.5	90	4	US-08-556-978B-93	Sequence 93, App	c 199	22.8	0.5	65	4	US-09-415-784-32	Sequence 32, Appl
127	23.4	0.5	90	5	PCT-US92-10983-131	Sequence 131, App	c 200	22.8	0.5	65	4	US-09-415-784-32	Sequence 32, Appl
128	23.4	0.5	93	4	US-09-042-353-248	Sequence 248, App	c 201	22.8	0.5	65	4	US-09-415-784-32	Sequence 32, Appl
129	23.4	0.5	93	4	US-08-758-417A-96	Sequence 96, App	c 202	22.8	0.5	65	4	US-09-415-784-32	Sequence 32, Appl
130	23.4	0.5	96	1	US-08-371-121-9	Sequence 96, App	c 203	22.8	0.5	76	1	US-07-988-430-81	Sequence 81, Appl
131	23.4	0.5	96	1	US-09-384-212-4	Sequence 9, Appl	c 204	22.8	0.5	76	1	US-08-425-336-79	Sequence 79, Appl
132	23.4	0.5	81	1	US-08-399-412A-93	Sequence 93, Appl	c 205	22.8	0.5	76	1	US-08-425-336-79	Sequence 79, Appl
133	23.2	0.5	81	2	US-08-169-948B-21	Sequence 21, Appl	c 206	22.8	0.5	76	1	US-08-425-336-79	Sequence 79, Appl
134	23.2	0.5	81	2	US-08-448-873-21	Sequence 21, Appl	c 207	22.8	0.5	76	2	US-08-646-366-79	Sequence 79, Appl
135	23.2	0.5	81	4	US-08-382-452D-21	Sequence 21, Appl	c 208	22.8	0.5	76	3	US-08-646-366-79	Sequence 79, Appl
136	23.2	0.5	83	1	US-08-175-155-34	Sequence 34, Appl	c 209	22.8	0.5	76	3	US-09-136-368-79	Sequence 79, Appl
137	23.2	0.5	83	1	US-08-477-509B-69	Sequence 69, Appl	c 210	22.8	0.5	76	4	US-09-610-838-79	Sequence 79, Appl
138	23.2	0.5	83	2	US-08-707-237A-40	Sequence 40, Appl	c 211	22.8	0.5	85	4	PCT-US92-09487-81	Sequence 81, Appl
139	23.2	0.5	83	3	US-08-482-085B-69	Sequence 69, Appl	c 212	22.8	0.5	85	4	US-09-220-510B-5	Sequence 5, Appl
140	23.2	0.5	83	4	US-09-444-791A-69	Sequence 69, Appl	c 213	22.6	0.4	63	1	US-08-447-169A-212	Sequence 212, App
141	23.2	0.5	90	4	US-09-461-697-434	Sequence 434, App	c 214	22.6	0.4	75	1	US-08-750-419A-4	Sequence 4, Appl
142	23.2	0.5	91	4	US-09-205-995-69	Sequence 69, App	c 215	22.6	0.4	75	1	US-08-318-193-72	Sequence 72, Appl
143	23.2	0.5	97	2	US-08-479-729-7	Sequence 7, Appl	c 216	22.6	0.4	78	6	5457037-7	Patent No. 5457037
144	23.2	0.5	97	5	PCT-US96-08089-7	Sequence 7, Appl	c 217	22.6	0.4	93	2	US-08-707-237A-100	Sequence 100, App
145	23.2	0.5	98	1	US-07-889-651-16	Sequence 16, Appl	c 218	22.6	0.4	93	3	US-08-642-246-22	Sequence 22, Appl
146	23.2	0.5	100	1	US-07-797-556-7	Sequence 7, Appl	c 219	22.6	0.4	93	4	US-08-806-029-20	Sequence 20, Appl
147	23	0.5	66	4	US-08-556-978B-42	Sequence 42, App	c 220	22.6	0.4	93	4	US-09-451-206-22	Sequence 22, Appl
148	23	0.5	68	1	US-08-180-195-26	Sequence 26, App	c 221	22.6	0.4	96	1	PCT-US96-06229-22	Sequence 22, Appl
149	23	0.5	68	1	US-08-477-329-26	Sequence 26, App	c 222	22.6	0.4	96	1	US-08-208-886C-60	Sequence 60, Appl
150	23	0.5	68	2	US-08-475-458-26	Sequence 26, App	c 223	22.6	0.4	96	1	US-08-208-886C-72	Sequence 72, Appl
151	23	0.5	68	3	US-08-980-400-26	Sequence 26, App	c 224	22.6	0.4	96	1	US-08-704-744-60	Sequence 60, Appl
152	23	0.5	68	4	US-09-583-459A-26	Sequence 26, App	c 225	22.6	0.4	96	1	US-08-704-744-72	Sequence 72, Appl
153	23	0.5	68	4	US-09-583-459A-26	Sequence 26, App	c 226	22.6	0.4	96	1	US-08-469-551-60	Sequence 60, Appl
154	23	0.5	68	4	US-09-583-459A-26	Sequence 26, App	c 227	22.6	0.4	96	2	US-08-290-799B-60	Sequence 60, Appl
155	23	0.5	68	4	US-09-583-459A-26	Sequence 26, App	c 228	22.6	0.4	97	3	US-08-825-852-38	Sequence 38, Appl
156	23	0.5	82	2	US-08-476-176B-24	Sequence 24, App	c 229	22.6	0.4	97	3	US-08-825-852-38	Sequence 38, Appl
157	23	0.5	82	3	US-08-127-721A-24	Sequence 24, App	c 230	22.6	0.4	97	4	US-09-052-888-44	Sequence 44, Appl
158	23	0.5	82	3	US-08-485-246A-24	Sequence 24, App	c 231	22.6	0.4	97	4	US-09-052-888-44	Sequence 44, Appl
159	23	0.5	89	2	US-08-481-658B-42	Sequence 42, App	c 232	22.6	0.4	100	1	US-08-208-886C-59	Sequence 59, Appl
160	23	0.5	89	2	US-08-477-504A-42	Sequence 42, App	c 233	22.6	0.4	100	1	US-08-208-886C-71	Sequence 71, Appl
161	23	0.5	89	2	US-08-486-756A-42	Sequence 42, App	c 234	22.6	0.4	100	1	US-08-704-744-59	Sequence 59, Appl
162	23	0.5	89	2	US-08-485-049D-42	Sequence 42, App	c 235	22.6	0.4	100	1	US-08-704-744-71	Sequence 71, Appl
163	23	0.5	89	3	US-08-787-739-42	Sequence 42, App	c 236	22.6	0.4	100	1	US-08-469-551-59	Sequence 59, Appl
164	23	0.5	89	3	US-08-487-077A-42	Sequence 42, App	c 237	22.4	0.4	100	2	US-08-290-799B-59	Sequence 59, Appl
165	23	0.5	89	3	US-08-485-863A-42	Sequence 42, App	c 238	22.4	0.4	60	5	PCT-US93-01901-27	Sequence 27, Appl
166	23	0.5	89	4	US-08-485-049D-42	Sequence 42, App	c 239	22.4	0.4	69	4	US-09-042-353-26	Sequence 26, App
167	23	0.5	89	4	US-09-178-115-42	Sequence 42, App	c 240	22.4	0.4	69	4	US-08-758-417A-94	Sequence 94, App
168	23	0.5	89	4	US-09-177-776-42	Sequence 42, App	c 241	22.4	0.4	72	1	US-07-826-928A-19	Sequence 19, Appl
169	23	0.5	96	1	US-08-447-422-11	Sequence 11, Appl	c 242	22.4	0.4	81	3	US-08-776-251-2	Sequence 2, Appl
170	23	0.5	96	1	US-08-755-587-4	Sequence 4, Appl	c 243	22.4	0.4	81	4	US-08-556-978B-90	Sequence 90, Appl
171	23	0.5	97	3	US-08-825-852-39	Sequence 39, Appl	c 244	22.4	0.4	81	4	US-09-497-933A-20	Sequence 20, Appl
172	23	0.5	97	3	US-08-825-852-40	Sequence 40, Appl	c 245	22.4	0.4	95	2	US-08-476-176B-20	Sequence 20, Appl
173	23	0.5	97	3	US-08-825-852-43	Sequence 43, Appl	c 246	22.4	0.4	95	3	US-08-127-721A-20	Sequence 20, Appl

247	22.4	0.4	96	1	US-08-447-422-12	Sequence 12, Appl	320	21.6	0.4	88	6	5194596-6	Patent No. 5194596
248	22.4	0.4	100	2	US-08-690-495-30	Sequence 30, Appl	321	21.6	0.4	94	4	US-09-025-769B-175	Sequence 175, Appl
249	22.4	0.4	100	2	US-08-690-494-30	Sequence 30, Appl	322	21.4	0.4	62	1	US-08-182-175A-98	Sequence 98, Appl
250	22.4	0.4	100	5	PCT-US95-07671-30	Sequence 30, Appl	323	21.4	0.4	62	1	US-08-446-908-15	Sequence 15, Appl
C 251	22.2	0.4	57	1	US-07-609-716-51	Sequence 51, Appl	324	21.4	0.4	62	1	US-08-231-205A-87	Sequence 87, Appl
C 252	22.2	0.4	57	1	US-07-609-716-52	Sequence 52, Appl	325	21.4	0.4	62	1	US-08-474-633A-87	Sequence 87, Appl
C 253	22.2	0.4	57	3	US-08-642-255-20	Sequence 20, Appl	326	21.4	0.4	62	2	US-08-871-161-15	Sequence 15, Appl
C 254	22.2	0.4	57	3	US-08-475-411A-51	Sequence 51, Appl	327	21.4	0.4	62	4	US-08-823-771-87	Sequence 87, Appl
C 255	22.2	0.4	57	3	US-08-475-411A-52	Sequence 52, Appl	328	21.4	0.4	62	5	PCT-US92-06412-98	Sequence 98, Appl
C 256	22.2	0.4	57	4	US-08-478-028A-51	Sequence 51, Appl	329	21.4	0.4	62	5	US-08-468-161-102	Sequence 102, Appl
C 257	22.2	0.4	57	4	US-08-478-028A-52	Sequence 52, Appl	330	21.4	0.4	69	3	US-09-273-685-102	Sequence 102, Appl
C 258	22.2	0.4	64	4	US-09-025-769B-110	Sequence 110, Appl	331	21.4	0.4	69	5	PCT-US95-11934-102	Sequence 5, Appl
C 259	22.2	0.4	67	2	US-07-977-284A-218	Sequence 218, Appl	332	21.4	0.4	75	1	US-08-237-716-5	Sequence 6, Appl
C 260	22.2	0.4	72	1	US-08-318-193-78	Sequence 78, Appl	333	21.4	0.4	75	1	US-08-237-716-6	Sequence 6, Appl
C 261	22.2	0.4	72	1	US-08-318-193-78	Sequence 78, Appl	334	21.4	0.4	75	3	US-08-131-567-2	Sequence 26, Appl
C 262	22.2	0.4	81	4	US-08-556-978B-91	Sequence 91, Appl	335	21.4	0.4	79	1	US-08-384-708A-226	Sequence 22, Appl
C 263	22.2	0.4	90	3	US-08-952-664-17	Sequence 17, Appl	336	21.4	0.4	79	2	US-08-470-939-22	Sequence 22, Appl
C 264	22.2	0.4	90	4	US-09-065-058-16	Sequence 16, Appl	337	21.4	0.4	79	4	US-08-687-421-318	Sequence 318, Appl
C 265	22.2	0.4	90	4	US-09-487-87A-17	Sequence 17, Appl	338	21.4	0.4	79	5	PCT-US96-09452-22	Sequence 22, Appl
C 266	22.2	0.4	91	4	US-09-404-879A-201	Sequence 201, Appl	339	21.4	0.4	85	2	US-08-852-807-8	Sequence 8, Appl
C 267	22.2	0.4	92	1	US-08-487-141B-109	Sequence 109, Appl	340	21.4	0.4	87	2	US-08-476-176B-22	Sequence 22, Appl
C 268	22.2	0.4	52	3	US-08-718-904-104	Sequence 104, Appl	341	21.4	0.4	87	3	US-08-127-721A-25	Sequence 25, Appl
C 269	22.2	0.4	55	1	US-07-826-928A-16	Sequence 16, Appl	342	21.4	0.4	87	3	US-08-485-246A-22	Sequence 22, Appl
C 270	22.2	0.4	66	1	US-08-180-195-23	Sequence 23, Appl	343	21.4	0.4	92	1	US-08-120-827-94	Sequence 94, Appl
C 271	22.2	0.4	66	1	US-08-477-329-23	Sequence 23, Appl	344	21.4	0.4	92	1	US-08-478-675-94	Sequence 94, Appl
C 272	22.2	0.4	66	2	US-08-475-458-23	Sequence 23, Appl	345	21.4	0.4	94	1	US-08-479-783A-59	Sequence 59, Appl
C 273	22.2	0.4	66	2	US-08-980-400-23	Sequence 23, Appl	346	21.4	0.4	94	1	US-08-479-725-59	Sequence 59, Appl
C 274	22.2	0.4	66	4	US-09-583-459A-23	Sequence 23, Appl	347	21.4	0.4	94	1	US-08-618-693-59	Sequence 59, Appl
C 275	22.2	0.4	66	4	US-09-583-210-23	Sequence 23, Appl	348	21.4	0.4	94	4	US-08-973-123A-48	Sequence 148, Appl
C 276	22.2	0.4	66	4	US-09-583-449A-23	Sequence 23, Appl	349	21.4	0.4	94	4	US-08-991-743C-59	Sequence 59, Appl
C 277	22.2	0.4	67	4	US-09-435-059-23	Sequence 23, Appl	350	21.4	0.4	94	5	PCT-US96-08014-148	Sequence 148, Appl
C 278	22.2	0.4	67	4	US-07-977-284A-245	Sequence 245, Appl	351	21.4	0.4	97	2	US-08-476-176B-25	Sequence 25, Appl
C 279	22.2	0.4	67	2	US-08-256-426B-245	Sequence 245, Appl	352	21.4	0.4	97	2	US-08-127-721A-25	Sequence 25, Appl
C 280	22.2	0.4	68	1	US-08-440-787A-52	Sequence 52, Appl	353	21.4	0.4	97	3	US-08-485-246A-25	Sequence 25, Appl
C 281	22.2	0.4	68	4	US-08-367-685-52	Sequence 52, Appl	354	21.4	0.4	97	3	US-08-825-852-11	Sequence 11, Appl
C 282	22.2	0.4	68	5	PCT-US91-07141-52	Sequence 52, Appl	355	21.4	0.4	97	3	US-08-825-852-12	Sequence 12, Appl
C 283	22.2	0.4	76	4	US-07-757-022B-3	Sequence 3, Appl	356	21.4	0.4	97	3	US-08-825-852-14	Sequence 14, Appl
C 284	22.2	0.4	81	1	US-08-170-290A-17	Sequence 17, Appl	357	21.4	0.4	97	4	US-09-052-888-11	Sequence 11, Appl
C 285	22.2	0.4	84	2	US-08-426-598B-10	Sequence 10, Appl	358	21.4	0.4	97	4	US-09-052-888-42	Sequence 42, Appl
C 286	22.2	0.4	89	2	US-08-716-308-4	Sequence 4, Appl	359	21.4	0.4	97	4	US-09-052-888-44	Sequence 44, Appl
C 287	22.2	0.4	92	4	US-09-506-729-48	Sequence 48, Appl	360	21.4	0.4	100	3	US-09-017-631-34	Sequence 34, Appl
C 288	22.2	0.4	95	4	US-09-511-625B-44	Sequence 44, Appl	361	21.4	0.4	100	4	US-09-018-138-12	Sequence 12, Appl
C 289	22.2	0.4	95	4	US-08-464-700-43	Sequence 43, Appl	362	21.2	0.4	60	4	US-09-284-627-14	Sequence 14, Appl
C 290	22.2	0.4	46	1	US-08-522-623-4	Sequence 4, Appl	363	21.2	0.4	64	2	US-08-874-678-13	Sequence 13, Appl
C 291	21.8	0.4	45	5	PCT-US93-03256-4	Sequence 4, Appl	364	21.2	0.4	64	3	US-08-643-839-43	Sequence 43, Appl
C 292	21.8	0.4	67	1	US-07-977-284A-258	Sequence 258, Appl	365	21.2	0.4	64	4	US-09-348-886-43	Sequence 43, Appl
C 293	21.8	0.4	67	2	US-08-256-426B-258	Sequence 258, Appl	366	21.2	0.4	69	1	US-08-215-138-3	Sequence 3, Appl
C 294	21.8	0.4	75	1	US-07-982-712-6	Sequence 6, Appl	367	21.2	0.4	73	4	US-08-407-344-3	Sequence 3, Appl
C 295	21.8	0.4	75	1	US-07-941-651-5	Sequence 5, Appl	368	21.2	0.4	73	4	US-09-303-069-20	Sequence 20, Appl
C 296	21.8	0.4	75	1	US-08-279-996-5	Sequence 5, Appl	369	21.2	0.4	73	4	US-09-134-250-20	Sequence 20, Appl
C 297	21.8	0.4	75	1	US-08-482-111-61	Sequence 61, Appl	370	21.2	0.4	82	2	US-08-481-658B-37	Sequence 37, Appl
C 298	21.8	0.4	83	1	US-08-175-155-35	Sequence 35, Appl	371	21.2	0.4	82	2	US-08-477-504A-37	Sequence 37, Appl
C 299	21.8	0.4	83	1	US-08-477-509B-70	Sequence 70, Appl	372	21.2	0.4	82	2	US-08-486-756A-37	Sequence 37, Appl
C 300	21.8	0.4	83	2	US-08-707-237A-41	Sequence 41, Appl	373	21.2	0.4	82	2	US-08-485-862B-37	Sequence 37, Appl
C 301	21.8	0.4	83	3	US-08-482-085B-70	Sequence 70, Appl	374	21.2	0.4	82	3	US-08-787-729-37	Sequence 37, Appl
C 302	21.8	0.4	83	4	US-09-444-791A-70	Sequence 70, Appl	375	21.2	0.4	82	3	US-08-487-077A-37	Sequence 37, Appl
C 303	21.8	0.4	86	2	US-08-870-518-12	Sequence 12, Appl	376	21.2	0.4	82	3	US-08-485-863A-37	Sequence 37, Appl
C 304	21.8	0.4	89	4	US-09-296-328A-2	Sequence 2, Appl	377	21.2	0.4	82	4	US-08-485-049D-37	Sequence 37, Appl
C 305	21.8	0.4	89	4	US-08-464-700-18	Sequence 18, Appl	378	21.2	0.4	82	4	US-09-177-776-37	Sequence 37, Appl
C 306	21.6	0.4	54	1	US-07-982-712-18	Sequence 18, Appl	379	21.2	0.4	82	4	US-09-177-776-37	Sequence 37, Appl
C 307	21.6	0.4	54	1	US-09-339-913B-52	Sequence 52, Appl	380	21.2	0.4	86	1	US-08-447-169A-218	Sequence 218, Appl
C 308	21.6	0.4	59	4	US-09-339-904A-52	Sequence 52, Appl	381	21.2	0.4	86	1	US-08-433-126A-180	Sequence 180, Appl
C 309	21.6	0.4	59	4	US-08-769-062B-52	Sequence 52, Appl	382	21.2	0.4	87	1	US-08-433-124A-180	Sequence 180, Appl
C 310	21.6	0.4	59	4	US-09-344-002B-52	Sequence 52, Appl	383	21.2	0.4	87	1	US-08-976-413A-180	Sequence 180, Appl
C 311	21.6	0.4	59	4	US-09-559-565C-14	Sequence 52, Appl	384	21.2	0.4	87	4	US-08-936-477-4	Sequence 4, Appl
C 312	21.6	0.4	60	3	US-08-654-623-14	Sequence 14, Appl	385	21.2	0.4	87	4	US-09-113-785-13	Sequence 13, Appl
C 313	21.6	0.4	64	4	US-09-025-769B-102	Sequence 102, Appl	386	21.2	0.4	87	5	PCT-US96-06059-180	Sequence 180, Appl
C 314	21.6	0.4	69	4	US-08-931-858B-209	Sequence 209, Appl	387	21.2	0.4	90	4	US-08-974-549A-667	Sequence 667, Appl
C 315	21.6	0.4	69	4	US-08-931-858B-210	Sequence 210, Appl	388	21.2	0.4	90	4	US-08-974-549A-668	Sequence 668, Appl
C 316	21.6	0.4	70	4	US-09-625-188-41	Sequence 41, Appl	389	21.2	0.4	93	2	US-08-553-501A-47	Sequence 47, Appl
C 317	21.6	0.4	72	4	US-09-404-879A-304	Sequence 304, Appl	390	21.2	0.4	93	3	US-09-205-231-47	Sequence 47, Appl
C 318	21.6	0.4	81	3	US-09-014-416-57	Sequence 57, Appl	391	21	0.4	29	4	US-09-657-042A-6	Sequence 6, Appl
C 319	21.6	0.4	84	4	US-08-943-731-204	Sequence 204, Appl	392	21	0.4	41	1	US-08-253-155A-82	Sequence 82, Appl

C 393	21	0.4	41	1	US-08-625-209A-11	Sequence 11, App1	466	21	0.4	93	5	PCT-US92-10983-133	Sequence 133, App
C 394	21	0.4	41	3	US-08-853-733B-11	Sequence 11, App1	C 467	21	0.4	94	1	US-07-626-618A-14	Sequence 14, App1
C 395	21	0.4	45	4	US-09-561-490E-35	Sequence 35, App1	C 468	21	0.4	94	1	US-07-928-611-14	Sequence 14, App1
C 396	21	0.4	47	4	US-09-641-638-877	Sequence 877, App	C 469	21	0.4	94	1	US-08-333-977-14	Sequence 14, App1
C 397	21	0.4	64	1	US-08-053-131-135	Sequence 135, App	C 470	21	0.4	94	1	US-08-053-131-144	Sequence 144, App
C 398	21	0.4	64	1	US-08-645-641-135	Sequence 135, App	C 471	21	0.4	94	1	US-08-645-641-144	Sequence 144, App
C 399	21	0.4	64	1	US-07-853-408B-135	Sequence 135, App	C 472	21	0.4	94	1	US-07-853-408B-144	Sequence 144, App
C 400	21	0.4	64	1	US-08-096-762-135	Sequence 135, App	C 473	21	0.4	94	1	US-08-096-762-144	Sequence 144, App
C 401	21	0.4	64	2	US-08-308-865-135	Sequence 135, App	C 474	21	0.4	94	2	US-08-308-865-144	Sequence 144, App
C 402	21	0.4	64	2	US-09-042-353-252	Sequence 252, App	C 475	21	0.4	94	2	US-08-487-812A-14	Sequence 14, App
C 403	21	0.4	64	4	US-08-758-417A-100	Sequence 100, App	C 476	21	0.4	94	4	US-09-060-694-14	Sequence 14, App1
C 404	21	0.4	64	5	PCT-US92-10983-135	Sequence 135, App	C 477	21	0.4	94	4	US-09-042-353-261	Sequence 261, App
C 405	21	0.4	65	4	US-09-142-355B-8	Sequence 8, App11	C 478	21	0.4	94	4	US-08-758-417A-109	Sequence 109, App
C 406	21	0.4	72	1	US-08-208-886C-56	Sequence 56, App1	C 479	21	0.4	94	4	US-09-378-074-14	Sequence 14, App1
C 407	21	0.4	72	1	US-08-704-744-56	Sequence 56, App1	C 480	21	0.4	94	5	PCT-US92-10983-144	Sequence 144, App
C 408	21	0.4	72	1	US-08-469-557-56	Sequence 56, App1	C 481	21	0.4	94	5	PCT-US93-07370-14	Sequence 14, App
C 409	21	0.4	72	2	US-08-290-793B-56	Sequence 56, App1	C 482	21	0.4	94	5	US-09-042-353-245	Sequence 245, App
C 410	21	0.4	73	2	US-09-010-928B-8	Sequence 8, App11	C 483	21	0.4	99	4	US-08-758-417A-93	Sequence 93, App1
C 411	21	0.4	74	3	US-08-789-333F-61	Sequence 61, App1	C 484	21	0.4	36	1	US-08-309-512-60	Sequence 60, App1
C 412	21	0.4	74	4	US-08-787-738B-61	Sequence 61, App1	C 485	20.8	0.4	42	4	US-09-250-609-34	Sequence 34, App1
C 413	21	0.4	76	1	US-08-208-886C-55	Sequence 55, App1	C 486	20.8	0.4	45	5	PCT-US95-14639-8	Sequence 8, App1
C 414	21	0.4	76	1	US-08-704-744-55	Sequence 55, App1	C 487	20.8	0.4	47	4	US-09-641-638-1224	Sequence 1224, App
C 415	21	0.4	76	1	US-08-469-557-55	Sequence 55, App1	C 488	20.8	0.4	54	1	US-08-015-180-8	Sequence 8, App1
C 416	21	0.4	76	2	US-08-290-793B-55	Sequence 55, App1	C 489	20.8	0.4	54	1	US-08-649-196-8	Sequence 8, App1
C 417	21	0.4	78	4	US-08-976-183A-1	Sequence 1, App1	C 490	20.8	0.4	60	3	US-08-983-607-14	Sequence 14, App1
C 418	21	0.4	78	4	US-08-976-183A-3	Sequence 3, App11	C 491	20.8	0.4	61	1	US-08-616-133-22	Sequence 22, App1
C 419	21	0.4	78	4	US-09-425-638A-36	Sequence 36, App1	C 492	20.8	0.4	61	1	US-08-802-985-22	Sequence 22, App1
C 420	21	0.4	78	4	US-09-543-004-36	Sequence 36, App1	C 493	20.8	0.4	65	1	US-08-616-133-21	Sequence 21, App1
C 421	21	0.4	80	1	US-08-208-886C-36	Sequence 36, App1	C 494	20.8	0.4	65	1	US-08-802-985-21	Sequence 21, App1
C 422	21	0.4	80	1	US-08-704-744-36	Sequence 36, App1	C 495	20.8	0.4	66	4	US-08-556-978B-41	Sequence 41, App1
C 423	21	0.4	80	1	US-08-469-557-36	Sequence 36, App1	C 496	20.8	0.4	67	1	US-07-977-284A-23	Sequence 23, App
C 424	21	0.4	80	2	US-08-290-793B-36	Sequence 36, App1	C 497	20.8	0.4	67	2	US-08-256-426B-23	Sequence 223, App
C 425	21	0.4	80	4	US-08-978-806-6	Sequence 6, App11	C 498	20.8	0.4	68	1	US-08-428-733A-50	Sequence 50, App1
C 426	21	0.4	81	1	US-08-238-863-45	Sequence 22, App1	C 499	20.8	0.4	72	1	US-07-977-284A-229	Sequence 229, App
C 427	21	0.4	81	1	US-08-443-407-22	Sequence 45, App1	C 500	20.8	0.4	72	2	US-08-256-426B-229	Sequence 229, App
C 428	21	0.4	81	1	US-08-443-407-45	Sequence 45, App1	C 501	20.8	0.4	73	3	US-07-854-845B-2	Sequence 2, App11
C 429	21	0.4	81	5	PCT-US95-05600-166	Sequence 166, App	C 502	20.8	0.4	77	3	US-08-666-354A-10	Sequence 10, App1
C 430	21	0.4	81	5	PCT-US95-05600-189	Sequence 189, App	C 503	20.8	0.4	80	1	US-08-290-373B-13	Sequence 13, App1
C 431	21	0.4	84	2	US-08-788-943A-14	Sequence 14, App1	C 504	20.8	0.4	81	4	US-08-556-978B-27	Sequence 27, App1
C 432	21	0.4	84	5	PCT-US96-00952-15	Sequence 15, App1	C 505	20.8	0.4	81	4	US-08-556-978B-28	Sequence 28, App1
C 433	21	0.4	86	1	US-08-433-126A-69	Sequence 69, App1	C 506	20.8	0.4	81	4	US-09-497-933A-20	Sequence 20, App1
C 434	21	0.4	86	1	US-08-433-126A-69	Sequence 69, App1	C 507	20.8	0.4	82	1	US-08-290-373B-14	Sequence 14, App1
C 435	21	0.4	86	1	US-08-433-126A-69	Sequence 69, App1	C 508	20.8	0.4	83	4	US-09-133-321-10	Sequence 10, App1
C 436	21	0.4	86	3	US-08-976-413A-69	Sequence 69, App1	C 509	20.8	0.4	86	4	US-08-974-578A-69	Sequence 69, App1
C 437	21	0.4	86	5	PCT-US96-06059-69	Sequence 69, App1	C 510	20.8	0.4	90	4	US-08-481-658B-32	Sequence 32, App1
C 438	21	0.4	87	5	US-08-788-943A-15	Sequence 15, App1	C 511	20.8	0.4	90	4	US-08-477-504A-32	Sequence 32, App1
C 439	21	0.4	87	2	US-08-449-287-19	Sequence 19, App1	C 512	20.8	0.4	90	4	US-08-556-978B-74	Sequence 74, App1
C 440	21	0.4	87	5	PCT-US96-00952-16	Sequence 16, App1	C 513	20.8	0.4	92	1	US-08-353-400-19	Sequence 19, App1
C 441	21	0.4	88	4	US-09-335-012-25	Sequence 25, App1	C 514	20.8	0.4	93	2	US-08-481-658B-32	Sequence 32, App1
C 442	21	0.4	88	4	US-09-941-964-25	Sequence 25, App1	C 515	20.8	0.4	93	2	US-08-486-756A-32	Sequence 32, App1
C 443	21	0.4	88	6	5219739-6	Patent No. 5219739	C 516	20.8	0.4	93	2	US-08-485-865B-32	Sequence 32, App1
C 444	21	0.4	90	4	US-08-464-700-41	Sequence 41, App1	C 517	20.8	0.4	93	2	US-08-485-865B-32	Sequence 32, App1
C 445	21	0.4	90	4	US-08-556-978B-74	Sequence 74, App1	C 518	20.8	0.4	93	3	US-08-787-739-32	Sequence 32, App1
C 446	21	0.4	90	4	US-09-480-251-10	Sequence 10, App1	C 519	20.8	0.4	93	3	US-08-487-077A-32	Sequence 32, App1
C 447	21	0.4	92	1	US-08-319-467-5	Sequence 5, App11	C 520	20.8	0.4	93	3	US-08-485-865A-32	Sequence 32, App1
C 448	21	0.4	92	2	US-08-781-084-5	Sequence 5, App11	C 521	20.8	0.4	93	3	US-08-485-865A-32	Sequence 32, App1
C 449	21	0.4	92	5	PCT-US95-12914-5	Sequence 5, App11	C 522	20.8	0.4	93	4	US-09-178-115-32	Sequence 32, App1
C 450	21	0.4	93	1	US-08-458-120-3	Sequence 3, App11	C 523	20.8	0.4	93	4	US-09-177-776-33	Sequence 32, App1
C 451	21	0.4	93	1	US-08-053-131-133	Sequence 133, App	C 524	20.8	0.4	95	2	US-08-596-387B-103	Sequence 103, App
C 452	21	0.4	93	1	US-08-353-400-20	Sequence 20, App1	C 525	20.8	0.4	95	4	US-09-067-615-103	Sequence 103, App
C 453	21	0.4	93	1	US-08-645-641-133	Sequence 133, App	C 526	20.8	0.4	95	5	PCT-US95-09816A-103	Sequence 103, App
C 454	21	0.4	93	1	US-07-853-408B-133	Sequence 133, App	C 527	20.8	0.4	98	1	US-08-088-658-42	Sequence 42, App1
C 455	21	0.4	93	1	US-08-096-762-133	Sequence 133, App	C 528	20.8	0.4	98	2	US-08-471-907A-42	Sequence 42, App1
C 456	21	0.4	93	2	US-08-308-865-133	Sequence 133, App	C 529	20.8	0.4	99	2	US-08-381-637-30	Sequence 30, App1
C 457	21	0.4	93	2	US-08-187-186A-3	Sequence 3, App11	C 530	20.8	0.4	100	1	US-08-129-930B-90	Sequence 90, App1
C 458	21	0.4	93	2	US-08-442-497C-3	Sequence 3, App11	C 531	20.8	0.4	100	3	US-08-431-081B-13	Sequence 13, App1
C 459	21	0.4	93	4	US-09-042-353-250	Sequence 250, App	C 532	20.8	0.4	100	3	US-08-608-016-13	Sequence 13, App1
C 460	21	0.4	93	4	US-08-556-978B-24	Sequence 24, App1	C 533	20.8	0.4	100	4	US-08-134-346A-45	Sequence 45, App1
C 461	21	0.4	93	4	US-08-556-978B-25	Sequence 25, App1	C 534	20.8	0.4	100	4	US-08-976-288A-90	Sequence 90, App1
C 462	21	0.4	93	4	US-08-556-978B-87	Sequence 87, App1	C 535	20.8	0.4	100	4	US-09-497-933A-13	Sequence 13, App1
C 463	21	0.4	93	4	US-08-556-978B-88	Sequence 88, App1	C 536	20.8	0.4	100	4	US-09-298-886-20	Sequence 20, App1
C 464	21	0.4	93	4	US-08-758-417A-98	Sequence 98, App1	C 537	20.6	0.4	27	3	US-08-594-455-58	Sequence 58, App1
C 465	21	0.4	93	4	US-09-333-033-3	Sequence 3, App11	C 538	20.6	0.4	27	3	US-09-258-408-58	Sequence 58, App1



C 539	20.6	0.4	31	2	US-08-560-916-3	Sequence 3, Appl1	612	20.6	0.4	96	3	US-08-836-561-72	Sequence 72, Appl1
C 540	20.6	0.4	31	2	US-08-676-841-3	Sequence 3, Appl1	613	20.6	0.4	96	3	US-08-484-322-6	Sequence 6, Appl1
C 541	20.6	0.4	51	1	US-08-068-747-1	Sequence 1, Appl1	C 614	20.6	0.4	100	1	US-08-145-705A-3	Sequence 1, Appl1
C 542	20.6	0.4	55	2	PCT-US93-01901-29	Sequence 29, Appl1	615	20.4	0.4	34	4	US-09-085-720-14	Sequence 14, Appl1
C 543	20.6	0.4	56	2	US-08-448-418-63	Sequence 62, Appl1	616	20.4	0.4	41	5	PCT-US93-01901-25	Sequence 25, Appl1
C 544	20.6	0.4	59	4	US-09-091-814-62	Sequence 63, Appl1	617	20.4	0.4	62	4	US-09-339-913B-51	Sequence 51, Appl1
C 545	20.6	0.4	61	3	US-09-023-228B-39	Sequence 39, Appl1	618	20.4	0.4	62	4	US-09-339-904A-51	Sequence 51, Appl1
C 546	20.6	0.4	61	3	US-09-163-025B-39	Sequence 39, Appl1	619	20.4	0.4	62	4	US-08-769-062B-51	Sequence 51, Appl1
C 547	20.6	0.4	66	4	US-09-091-814-61	Sequence 61, Appl1	620	20.4	0.4	62	4	US-09-744-002B-51	Sequence 51, Appl1
C 548	20.6	0.4	69	3	US-08-329-799-39	Sequence 39, Appl1	621	20.4	0.4	62	4	US-09-559-565C-51	Sequence 51, Appl1
C 549	20.6	0.4	70	4	US-09-364-380-22	Sequence 22, Appl1	622	20.4	0.4	63	5	PCT-US91-03680-8	Sequence 8, Appl1
C 550	20.6	0.4	72	1	US-07-832-905B-11	Sequence 11, Appl1	623	20.4	0.4	64	2	US-07-814-220-22	Sequence 22, Appl1
C 551	20.6	0.4	72	2	US-08-700-757-11	Sequence 11, Appl1	624	20.4	0.4	64	2	US-07-812-421-23	Sequence 23, Appl1
C 552	20.6	0.4	72	4	US-08-463-691-11	Sequence 11, Appl1	625	20.4	0.4	68	1	US-07-977-284A-234	Sequence 234, Appl1
C 553	20.6	0.4	72	4	US-08-255-235-15	Sequence 15, Appl1	626	20.4	0.4	68	2	US-08-256-426B-234	Sequence 234, Appl1
C 554	20.6	0.4	72	4	US-09-123-728-31	Sequence 31, Appl1	627	20.4	0.4	69	4	US-09-025-769B-153	Sequence 153, Appl1
C 555	20.6	0.4	74	4	US-08-589-109A-4	Sequence 4, Appl1	628	20.4	0.4	70	2	US-08-488-402A-4	Sequence 4, Appl1
C 556	20.6	0.4	75	3	US-08-718-904-93	Sequence 93, Appl1	629	20.4	0.4	70	2	US-08-484-552A-4	Sequence 4, Appl1
C 557	20.6	0.4	75	4	US-08-556-978B-53	Sequence 53, Appl1	630	20.4	0.4	70	5	PCT-US96-09472-4	Sequence 4, Appl1
C 558	20.6	0.4	75	4	US-08-556-978B-54	Sequence 54, Appl1	631	20.4	0.4	71	2	US-08-488-402A-50	Sequence 50, Appl1
C 559	20.6	0.4	77	2	US-08-785-750-10	Sequence 10, Appl1	632	20.4	0.4	71	2	US-08-484-552A-50	Sequence 50, Appl1
C 560	20.6	0.4	77	4	US-09-121-162-4	Sequence 4, Appl1	633	20.4	0.4	71	5	PCT-US96-09472-50	Sequence 50, Appl1
C 561	20.6	0.4	77	4	US-09-205-337-10	Sequence 10, Appl1	634	20.4	0.4	72	2	US-07-814-220-23	Sequence 23, Appl1
C 562	20.6	0.4	81	1	US-08-411-795B-149	Sequence 149, Appl1	635	20.4	0.4	72	2	US-07-812-421-23	Sequence 23, Appl1
C 563	20.6	0.4	81	1	US-08-411-795B-150	Sequence 149, Appl1	636	20.4	0.4	74	3	US-09-173-941-116	Sequence 116, Appl1
C 564	20.6	0.4	81	1	US-08-411-796-149	Sequence 149, Appl1	637	20.4	0.4	77	1	US-08-430-709-38	Sequence 38, Appl1
C 565	20.6	0.4	81	1	US-08-411-796-150	Sequence 150, Appl1	638	20.4	0.4	77	2	US-07-829-461A-19	Sequence 19, Appl1
C 566	20.6	0.4	81	1	US-08-469-319A-149	Sequence 149, Appl1	639	20.4	0.4	77	2	US-08-918-304A-38	Sequence 38, Appl1
C 567	20.6	0.4	81	2	US-08-469-319A-150	Sequence 150, Appl1	C 640	20.4	0.4	77	4	US-09-258-797-91	Sequence 91, Appl1
C 568	20.6	0.4	81	2	US-07-989-847-23	Sequence 23, Appl1	641	20.4	0.4	77	4	US-09-407-234-38	Sequence 38, Appl1
C 569	20.6	0.4	81	3	US-08-471-039-149	Sequence 149, Appl1	642	20.4	0.4	77	4	US-09-197-649-20	Sequence 20, Appl1
C 570	20.6	0.4	81	3	US-08-471-039-150	Sequence 150, Appl1	643	20.4	0.4	81	1	US-07-958-222A-4	Sequence 4, Appl1
C 571	20.6	0.4	81	4	US-08-469-411-23	Sequence 23, Appl1	C 644	20.4	0.4	84	4	US-09-461-697-63	Sequence 63, Appl1
C 572	20.6	0.4	81	4	US-08-764-114-149	Sequence 149, Appl1	645	20.4	0.4	85	3	US-08-448-194-47	Sequence 47, Appl1
C 573	20.6	0.4	81	4	US-08-764-114-150	Sequence 150, Appl1	646	20.4	0.4	85	4	US-08-867-921-47	Sequence 47, Appl1
C 574	20.6	0.4	81	4	US-08-469-419-149	Sequence 149, Appl1	C 647	20.4	0.4	88	2	US-08-483-528B-85	Sequence 85, Appl1
C 575	20.6	0.4	81	4	US-08-469-419-150	Sequence 150, Appl1	C 648	20.4	0.4	88	2	US-08-860-882A-59	Sequence 59, Appl1
C 576	20.6	0.4	81	5	PCT-US93-11198-149	Sequence 149, Appl1	C 649	20.4	0.4	88	3	US-08-673-799C-85	Sequence 85, Appl1
C 577	20.6	0.4	81	5	PCT-US93-11198-150	Sequence 150, Appl1	C 650	20.4	0.4	88	4	US-09-171-945-118	Sequence 118, Appl1
C 578	20.6	0.4	84	1	US-07-808-451-3	Sequence 3, Appl1	C 651	20.4	0.4	88	4	US-09-393-385B-85	Sequence 85, Appl1
C 579	20.6	0.4	84	1	US-07-808-451-4	Sequence 4, Appl1	C 652	20.4	0.4	88	4	US-09-011-769A-41	Sequence 41, Appl1
C 580	20.6	0.4	84	1	US-07-989-845-19	Sequence 19, Appl1	C 653	20.4	0.4	88	4	US-08-448-194-48	Sequence 48, Appl1
C 581	20.6	0.4	84	1	US-07-989-844-3	Sequence 3, Appl1	C 654	20.4	0.4	89	4	US-08-867-921-48	Sequence 48, Appl1
C 582	20.6	0.4	84	1	US-08-161-044-3	Sequence 3, Appl1	C 655	20.4	0.4	89	4	US-08-530-569B-14	Sequence 14, Appl1
C 583	20.6	0.4	84	1	US-08-240-121-3	Sequence 3, Appl1	656	20.4	0.4	90	4	US-08-974-549A-695	Sequence 695, Appl1
C 584	20.6	0.4	84	1	US-08-451-241-3	Sequence 3, Appl1	C 657	20.4	0.4	90	4	US-08-974-549A-696	Sequence 696, Appl1
C 585	20.6	0.4	84	1	US-08-443-568B-34	Sequence 34, Appl1	658	20.4	0.4	91	6	5194595-7	Patent No. 5194595
C 586	20.6	0.4	84	1	US-08-443-568B-35	Sequence 35, Appl1	659	20.4	0.4	93	3	US-08-448-194-49	Sequence 49, Appl1
C 587	20.6	0.4	84	5	PCT-US92-10625-3	Sequence 3, Appl1	660	20.4	0.4	93	4	US-08-867-921-49	Sequence 49, Appl1
C 588	20.6	0.4	84	5	PCT-US92-10625-4	Sequence 4, Appl1	661	20.4	0.4	93	5	PCT-US96-01600-16	Sequence 16, Appl1
C 589	20.6	0.4	84	5	PCT-US93-11297-3	Sequence 3, Appl1	662	20.4	0.4	95	1	US-08-479-783A-75	Sequence 75, Appl1
C 590	20.6	0.4	84	5	PCT-US93-11298-19	Sequence 19, Appl1	663	20.4	0.4	95	1	US-08-479-725-75	Sequence 75, Appl1
C 591	20.6	0.4	84	5	PCT-US94-06997-34	Sequence 34, Appl1	664	20.4	0.4	95	1	US-08-618-653-75	Sequence 75, Appl1
C 592	20.6	0.4	84	5	PCT-US94-06997-35	Sequence 35, Appl1	665	20.4	0.4	95	4	US-08-973-124-164	Sequence 164, Appl1
C 593	20.6	0.4	86	4	US-08-836-261A-10	Sequence 10, Appl1	666	20.4	0.4	95	4	US-08-991-743C-75	Sequence 75, Appl1
C 594	20.6	0.4	87	4	US-08-729-601A-16	Sequence 16, Appl1	667	20.4	0.4	95	5	PCT-US96-08014-164	Sequence 164, Appl1
C 595	20.6	0.4	88	4	US-08-686-968C-126	Sequence 126, Appl1	668	20.4	0.4	97	1	US-08-454-455-1	Sequence 1, Appl1
C 596	20.6	0.4	88	4	US-08-686-968C-133	Sequence 133, Appl1	669	20.4	0.4	99	3	US-08-721-979A-36	Sequence 36, Appl1
C 597	20.6	0.4	88	4	US-08-686-968C-138	Sequence 138, Appl1	670	20.4	0.4	99	3	US-08-836-501-36	Sequence 36, Appl1
C 598	20.6	0.4	88	4	US-08-686-968C-152	Sequence 152, Appl1	671	20.4	0.4	99	4	US-09-654-289-36	Sequence 36, Appl1
C 599	20.6	0.4	88	4	US-08-686-968C-155	Sequence 155, Appl1	C 672	20.4	0.4	100	1	US-08-145-705A-27	Sequence 27, Appl1
C 600	20.6	0.4	89	1	US-08-229-515A-5	Sequence 5, Appl1	C 673	20.4	0.4	100	1	US-08-655-086-3	Sequence 3, Appl1
C 601	20.6	0.4	89	1	US-08-645-865-5	Sequence 5, Appl1	C 674	20.2	0.4	40	4	US-08-991-840A-15	Sequence 15, Appl1
C 602	20.6	0.4	91	1	US-08-317-403A-39	Sequence 39, Appl1	675	20.2	0.4	47	4	US-09-641-638B-876	Sequence 876, Appl1
C 603	20.6	0.4	91	1	US-08-471-985A-39	Sequence 39, Appl1	676	20.2	0.4	47	4	US-09-641-638B-876	Sequence 876, Appl1
C 604	20.6	0.4	91	4	US-09-328-750A-5	Sequence 39, Appl1	677	20.2	0.4	51	4	US-09-401-171C-9	Sequence 9, Appl1
C 605	20.6	0.4	91	5	PCT-US95-12401A-39	Sequence 39, Appl1	C 678	20.2	0.4	57	2	US-08-225-224-9	Sequence 9, Appl1
C 606	20.6	0.4	96	1	US-08-055-917-2	Sequence 2, Appl1	C 679	20.2	0.4	57	2	US-09-018-760-1	Sequence 1, Appl1
C 607	20.6	0.4	96	1	US-08-095-068-2	Sequence 2, Appl1	C 680	20.2	0.4	57	3	US-08-722-258-9	Sequence 9, Appl1
C 608	20.6	0.4	96	1	US-08-140-721A-2	Sequence 2, Appl1	C 681	20.2	0.4	57	3	US-08-993-380-1	Sequence 1, Appl1
C 609	20.6	0.4	96	1	US-08-619-790C-3	Sequence 3, Appl1	C 682	20.2	0.4	57	3	PCT-US95-04468-9	Sequence 9, Appl1
C 610	20.6	0.4	96	2	US-07-785-565A-2	Sequence 2, Appl1	C 683	20.2	0.4	59	3	US-08-918-406-4	Sequence 4, Appl1
C 611	20.6	0.4	96	3	US-08-836-561-60	Sequence 60, Appl1	C 684	20.2	0.4	60	4	US-09-018-138B-23	Sequence 23, Appl1

C 685	20.2	0.4	63	2	US-08-658-322-10	Sequence 10, Appl
C 686	20.2	0.4	67	1	US-08-318-193-73	Sequence 73, Appl
C 687	20.2	0.4	69	1	US-08-215-138-20	Sequence 20, Appl
C 688	20.2	0.4	69	1	US-08-407-344-20	Sequence 20, Appl
C 689	20.2	0.4	69	3	US-08-480-173A-32	Sequence 32, Appl
C 690	20.2	0.4	69	3	US-08-484-408A-32	Sequence 32, Appl
C 691	20.2	0.4	70	1	US-07-982-712-23	Sequence 23, Appl
C 692	20.2	0.4	72	1	US-08-252-057-36	Sequence 26, Appl
C 693	20.2	0.4	74	2	US-08-788-943A-31	Sequence 31, Appl
C 694	20.2	0.4	74	4	US-09-341-007B-21	Sequence 21, Appl
C 695	20.2	0.4	75	2	US-08-788-943A-34	Sequence 34, Appl
C 696	20.2	0.4	75	2	US-08-388-353-648	Sequence 648, App
C 697	20.2	0.4	78	3	US-08-488-551B-648	Sequence 648, App
C 698	20.2	0.4	78	3	US-09-928-881-22	Sequence 22, Appl
C 699	20.2	0.4	78	3	US-09-543-921-22	Sequence 22, Appl
C 700	20.2	0.4	80	1	US-07-796-106-11	Sequence 11, Appl
C 701	20.2	0.4	80	1	US-08-458-023B-15	Sequence 15, Appl
C 702	20.2	0.4	81	3	US-08-651-136C-73	Sequence 73, Appl
C 703	20.2	0.4	81	3	US-09-173-941-117	Sequence 117, App
C 704	20.2	0.4	81	3	US-09-229-911A-73	Sequence 73, Appl
C 705	20.2	0.4	83	1	US-08-076-090-5	Sequence 5, Appl1
C 706	20.2	0.4	83	5	PCT-US94-06661-5	Sequence 5, Appl1
C 707	20.2	0.4	85	4	US-09-364-543-40	Sequence 40, Appl1
C 708	20.2	0.4	86	1	US-08-471-033-44	Sequence 44, Appl
C 709	20.2	0.4	86	2	US-08-471-044-44	Sequence 44, Appl
C 710	20.2	0.4	86	2	US-08-463-483A-44	Sequence 44, Appl
C 711	20.2	0.4	86	2	US-08-471-046A-44	Sequence 44, Appl
C 712	20.2	0.4	86	2	US-08-470-566B-44	Sequence 44, Appl
C 713	20.2	0.4	86	2	US-08-469-334-44	Sequence 44, Appl
C 714	20.2	0.4	86	3	US-09-300-529-44	Sequence 44, Appl
C 715	20.2	0.4	86	4	US-09-335-012-35	Sequence 35, Appl
C 716	20.2	0.4	86	4	US-09-941-964-35	Sequence 35, Appl
C 717	20.2	0.4	87	1	US-08-433-126A-157	Sequence 157, App
C 718	20.2	0.4	87	1	US-08-433-124A-157	Sequence 157, App
C 719	20.2	0.4	87	3	US-08-976-413A-157	Sequence 157, App
C 720	20.2	0.4	87	5	PCT-US96-06059-157	Sequence 157, App
C 721	20.2	0.4	88	1	US-07-796-106-2	Sequence 2, Appl1
C 722	20.2	0.4	88	1	US-07-796-106-12	Sequence 12, Appl
C 723	20.2	0.4	88	2	US-08-463-528B-85	Sequence 85, Appl
C 724	20.2	0.4	88	3	US-08-673-799C-85	Sequence 85, Appl
C 725	20.2	0.4	88	4	US-09-393-385B-85	Sequence 85, Appl
C 726	20.2	0.4	90	1	US-08-252-057-21	Sequence 21, Appl
C 727	20.2	0.4	90	1	US-08-184-751-21	Sequence 21, Appl
C 728	20.2	0.4	90	1	US-08-137-117D-104	Sequence 104, App
C 729	20.2	0.4	90	1	US-08-436-717-104	Sequence 104, App
C 730	20.2	0.4	90	4	US-07-827-691A-4	Sequence 4, Appl1
C 731	20.2	0.4	90	4	US-08-556-978B-73	Sequence 73, Appl
C 732	20.2	0.4	91	5	PCT-US94-08052-11	Sequence 11, Appl
C 733	20.2	0.4	92	2	US-08-392-771-7	Sequence 7, Appl1
C 734	20.2	0.4	97	3	US-08-825-852-39	Sequence 39, Appl
C 735	20.2	0.4	97	3	US-09-052-888-39	Sequence 39, Appl
C 736	20.2	0.4	99	1	US-08-208-886C-75	Sequence 75, Appl
C 737	20.2	0.4	99	1	US-08-208-886C-75	Sequence 76, Appl
C 738	20.2	0.4	99	1	US-08-378-761A-41	Sequence 41, Appl
C 739	20.2	0.4	99	1	US-08-485-286-41	Sequence 41, Appl
C 740	20.2	0.4	99	1	US-08-704-744-75	Sequence 75, Appl
C 741	20.2	0.4	99	1	US-08-704-744-76	Sequence 76, Appl
C 742	20.2	0.4	99	2	US-08-483-636-61	Sequence 61, Appl
C 743	20.2	0.4	99	2	US-08-483-632-61	Sequence 61, Appl
C 744	20.2	0.4	99	3	US-09-070-842A-2	Sequence 2, Appl1
C 745	20.2	0.4	99	3	US-08-859-998-500	Sequence 500, App
C 746	20.2	0.4	99	3	US-09-225-928-500	Sequence 500, App
C 747	20.2	0.4	35	4	US-09-297-269-10	Sequence 10, Appl
C 748	20.2	0.4	36	4	US-08-653-648A-55	Sequence 55, Appl
C 749	20.2	0.4	36	4	US-09-297-269-8	Sequence 8, Appl1
C 750	20.2	0.4	40	4	US-08-305-764C-42	Sequence 42, Appl
C 751	20.2	0.4	50	4	US-09-462-645C-35	Sequence 35, Appl
C 752	20.2	0.4	51	3	US-08-951-923-48	Sequence 48, Appl
C 753	20.2	0.4	53	1	US-08-429-181-57	Sequence 57, Appl
C 754	20.2	0.4	53	1	US-08-164-388-57	Sequence 57, Appl
C 755	20.2	0.4	56	1	US-08-211-202-53	Sequence 53, Appl
C 756	20.2	0.4	56	1	US-08-211-202-54	Sequence 54, Appl
C 757	20.2	0.4	56	1	US-08-211-202-58	Sequence 58, Appl
C 758	20.2	0.4	56	1	US-08-090-193-22	Sequence 22, Appl
C 759	20.2	0.4	56	2	US-08-488-031-22	Sequence 22, Appl
C 760	20.2	0.4	56	2	US-08-486-569-22	Sequence 22, Appl
C 761	20.2	0.4	56	2	US-08-488-027-22	Sequence 22, Appl
C 762	20.2	0.4	56	2	US-08-090-192-22	Sequence 22, Appl
C 763	20.2	0.4	56	2	US-08-482-663-22	Sequence 22, Appl
C 764	20.2	0.4	56	3	US-08-482-658-22	Sequence 22, Appl
C 765	20.2	0.4	56	3	US-08-470-349-22	Sequence 22, Appl
C 766	20.2	0.4	56	3	US-08-475-610-22	Sequence 22, Appl
C 767	20.2	0.4	56	3	PCT-US92-00277-22	Sequence 22, Appl
C 768	20.2	0.4	56	5	PCT-US92-00278-22	Sequence 22, Appl
C 769	20.2	0.4	57	1	US-07-916-034-18	Sequence 18, Appl
C 770	20.2	0.4	57	2	US-08-290-731C-8	Sequence 8, Appl1
C 771	20.2	0.4	58	1	US-07-972-033-63	Sequence 63, Appl
C 772	20.2	0.4	58	1	US-08-642-255-75	Sequence 75, Appl
C 773	20.2	0.4	58	2	US-08-448-418-64	Sequence 64, Appl
C 774	20.2	0.4	58	4	US-09-502-558-12	Sequence 12, Appl
C 775	20.2	0.4	60	4	US-09-339-913B-42	Sequence 42, Appl
C 776	20.2	0.4	60	4	US-09-339-904A-42	Sequence 42, Appl
C 777	20.2	0.4	60	4	US-08-769-062B-42	Sequence 42, Appl
C 778	20.2	0.4	60	4	US-09-344-002B-42	Sequence 42, Appl
C 779	20.2	0.4	60	4	US-09-559-565C-42	Sequence 42, Appl
C 780	20.2	0.4	62	1	US-08-182-175A-99	Sequence 99, Appl
C 781	20.2	0.4	62	1	US-08-474-633A-88	Sequence 88, Appl
C 782	20.2	0.4	62	4	US-08-823-771-88	Sequence 88, Appl
C 783	20.2	0.4	62	5	PCT-US92-06412-99	Sequence 99, Appl
C 784	20.2	0.4	66	4	US-09-291-874-20	Sequence 20, Appl
C 785	20.2	0.4	66	4	US-09-523-655-36	Sequence 36, Appl
C 786	20.2	0.4	66	4	US-08-790-963-87	Sequence 87, Appl
C 787	20.2	0.4	68	2	US-09-371-774-87	Sequence 87, Appl
C 788	20.2	0.4	69	4	US-08-931-858E-208	Sequence 208, App
C 789	20.2	0.4	69	4	US-08-931-858E-208	Sequence 208, App
C 790	20.2	0.4	70	1	US-08-487-141B-107	Sequence 107, App
C 791	20.2	0.4	70	2	US-08-927-561-107	Sequence 107, App
C 792	20.2	0.4	70	3	PCT-US96-09388-107	Sequence 107, App
C 793	20.2	0.4	71	1	US-08-453-742-3	Sequence 3, Appl1
C 794	20.2	0.4	71	1	US-08-454-464-3	Sequence 3, Appl1
C 795	20.2	0.4	71	1	US-08-453-222-3	Sequence 3, Appl1
C 796	20.2	0.4	71	1	US-08-452-802-3	Sequence 3, Appl1
C 797	20.2	0.4	71	1	US-08-453-956-3	Sequence 3, Appl1
C 798	20.2	0.4	71	1	US-08-086-631-3	Sequence 3, Appl1
C 799	20.2	0.4	71	2	US-08-452-930-3	Sequence 3, Appl1
C 800	20.2	0.4	71	3	US-08-831-132-17	Sequence 17, Appl
C 801	20.2	0.4	71	4	US-09-416-150-17	Sequence 17, Appl
C 802	20.2	0.4	71	5	PCT-US93-08143-3	Sequence 3, Appl1
C 803	20.2	0.4	72	2	US-07-814-220-23	Sequence 23, Appl
C 804	20.2	0.4	72	2	US-07-812-421-23	Sequence 23, Appl
C 805	20.2	0.4	74	3	US-08-718-904-97	Sequence 97, Appl
C 806	20.2	0.4	75	1	US-08-242-098-29	Sequence 29, Appl
C 807	20.2	0.4	75	2	US-07-916-098A-19	Sequence 19, Appl
C 808	20.2	0.4	75	2	US-07-916-098A-20	Sequence 20, Appl
C 809	20.2	0.4	78	4	US-09-043-303-12	Sequence 12, Appl
C 810	20.2	0.4	79	4	US-08-723-601A-27	Sequence 27, Appl
C 811	20.2	0.4	80	1	US-07-982-712-28	Sequence 28, Appl
C 812	20.2	0.4	81	1	US-07-744-282C-30	Sequence 30, Appl
C 813	20.2	0.4	81	4	US-08-556-978B-70	Sequence 70, Appl
C 814	20.2	0.4	81	4	US-08-556-978B-71	Sequence 71, Appl
C 815	20.2	0.4	81	5	PCT-US92-06921A-36	Sequence 36, Appl
C 816	20.2	0.4	83	1	US-08-420-443-5	Sequence 5, Appl1
C 817	20.2	0.4	84	1	US-08-039-778B-10	Sequence 10, Appl
C 818	20.2	0.4	86	4	US-09-565-596-17	Sequence 17, Appl
C 819	20.2	0.4	86	4	US-09-051-966-1	Sequence 1, Appl1
C 820	20.2	0.4	86	4	US-09-051-966-2	Sequence 2, Appl1
C 821	20.2	0.4	90	1	US-08-247-475-45	Sequence 45, Appl
C 822	20.2	0.4	90	1	US-08-479-650-45	Sequence 45, Appl
C 823	20.2	0.4	90	1	US-08-191-866D-70	Sequence 70, Appl
C 824	20.2	0.4	90	1	US-08-674-169-45	Sequence 45, Appl
C 825	20.2	0.4	90	2	US-08-185-949B-70	Sequence 70, Appl
C 826	20.2	0.4	90	4	US-08-974-549A-644	Sequence 644, App
C 827	20.2	0.4	91	3	US-08-464-700-16	Sequence 16, App
C 828	20.2	0.4	91	3	US-08-825-852-37	Sequence 37, Appl
C 829	20.2	0.4	91	4	US-09-053-868-37	Sequence 37, Appl
C 830	20.2	0.4	92	2	US-08-788-943A-23	Sequence 23, Appl

831	20	0.4	92	3	US-08-976-413A-419	Sequence 419, App	c 904	19.8	0.4	69	2	US-08-488-161-102	Sequence 102, App
832	20	0.4	93	2	US-08-381-637-28	Sequence 28, App	c 905	19.8	0.4	69	3	US-09-273-685-102	Sequence 102, App
833	20	0.4	93	3	US-08-284-516C-45	Sequence 45, App	c 906	19.8	0.4	69	3	US-08-480-117A-31	Sequence 31, App
834	20	0.4	93	4	US-09-537-911A-45	Sequence 45, App	c 907	19.8	0.4	69	3	US-08-484-408A-31	Sequence 31, App
835	20	0.4	95	2	US-08-525-742-29	Sequence 29, App	c 908	19.8	0.4	69	5	PCT-US95-11934-102	Sequence 102, App
836	20	0.4	96	3	US-08-484-322-20	Sequence 20, App	c 909	19.8	0.4	70	1	US-07-982-712-33	Sequence 33, App
837	20	0.4	97	2	US-08-525-742-27	Sequence 27, App	c 910	19.8	0.4	70	1	US-08-305-699-11	Sequence 11, App
838	20	0.4	98	2	US-08-263-911-21	Sequence 21, App	c 911	19.8	0.4	72	2	US-08-832-466-1	Sequence 1, App
839	20	0.4	98	4	US-08-991-789A-282	Sequence 282, App	c 912	19.8	0.4	72	2	US-08-822-028-56	Sequence 56, App
840	20	0.4	98	4	US-09-062-431-282	Sequence 282, App	c 913	19.8	0.4	72	2	US-08-479-285-56	Sequence 56, App
841	20	0.4	99	2	US-08-611-757-77	Sequence 77, App	c 914	19.8	0.4	72	6	5244792-14	Patent No. 5244792
842	20	0.4	99	3	US-08-721-979A-29	Sequence 29, App	c 915	19.8	0.4	73	1	US-08-453-104-15	Sequence 15, App
843	20	0.4	99	3	US-08-836-501-29	Sequence 29, App	c 916	19.8	0.4	73	1	US-08-175-155-41	Sequence 41, App
844	20	0.4	99	4	US-09-240-078-41	Sequence 41, App	c 917	19.8	0.4	73	1	US-08-477-509B-16	Sequence 16, App
845	20	0.4	99	4	US-09-654-289-29	Sequence 29, App	c 918	19.8	0.4	73	2	US-08-694-824-15	Sequence 15, App
846	20	0.4	99	5	PCT-US95-05980-77	Sequence 77, App	c 919	19.8	0.4	73	3	US-08-482-085B-76	Sequence 76, App
847	20	0.4	100	4	US-08-974-549A-643	Sequence 643, App	c 920	19.8	0.4	73	4	US-09-444-791A-76	Sequence 76, App
848	19.8	0.4	36	2	US-08-292-620A-1476	Sequence 1476, App	c 921	19.8	0.4	74	1	US-08-260-515-4	Sequence 4, App
849	19.8	0.4	36	3	US-08-585-684B-301	Sequence 301, App	c 922	19.8	0.4	77	1	US-08-384-708A-189	Sequence 189, App
850	19.8	0.4	36	3	US-09-071-845-1476	Sequence 1476, App	c 923	19.8	0.4	77	1	US-08-318-193-62	Sequence 62, App
851	19.8	0.4	36	4	US-09-038-073-301	Sequence 301, App	c 924	19.8	0.4	77	4	US-08-687-421-281	Sequence 281, App
852	19.8	0.4	42	4	US-09-162-484-7	Sequence 7, App	c 925	19.8	0.4	78	1	US-08-053-131-138	Sequence 138, App
853	19.8	0.4	42	4	US-09-641-638-1223	Sequence 1223, App	c 926	19.8	0.4	78	1	US-08-645-641-138	Sequence 138, App
854	19.8	0.4	50	1	US-08-445-640-20	Sequence 20, App	c 927	19.8	0.4	78	1	US-07-853-408B-138	Sequence 138, App
855	19.8	0.4	50	3	US-08-170-558-20	Sequence 20, App	c 928	19.8	0.4	78	1	US-08-096-762-138	Sequence 138, App
856	19.8	0.4	50	3	US-08-447-314-20	Sequence 20, App	c 929	19.8	0.4	78	2	US-08-308-865-138	Sequence 138, App
857	19.8	0.4	50	3	US-08-445-461-20	Sequence 20, App	c 930	19.8	0.4	78	4	US-09-042-353-255	Sequence 255, App
858	19.8	0.4	51	3	US-08-356-965-22	Sequence 22, App	c 931	19.8	0.4	78	4	US-08-758-417A-103	Sequence 103, App
859	19.8	0.4	51	3	US-08-621-700-22	Sequence 22, App	c 932	19.8	0.4	78	4	US-08-979-608A-31	Sequence 31, App
860	19.8	0.4	51	3	US-08-929-940-22	Sequence 22, App	c 933	19.8	0.4	78	5	PCT-US92-10963-138	Sequence 138, App
861	19.8	0.4	51	5	PCT-US95-03940-22	Sequence 22, App	c 934	19.8	0.4	80	3	US-08-933-983-42	Sequence 42, App
862	19.8	0.4	54	1	US-07-977-696C-51	Sequence 51, App	c 935	19.8	0.4	80	4	US-09-171-948-82	Sequence 82, App
863	19.8	0.4	54	1	US-08-129-930B-51	Sequence 51, App	c 936	19.8	0.4	80	4	US-09-459-427-22	Sequence 22, App
864	19.8	0.4	54	4	US-08-976-288A-51	Sequence 51, App	c 937	19.8	0.4	80	4	US-09-459-427-22	Sequence 22, App
865	19.8	0.4	55	4	US-09-007-005-12	Sequence 12, App	c 938	19.8	0.4	83	4	US-09-060-756-525	Sequence 525, App
866	19.8	0.4	55	4	US-09-244-796-12	Sequence 12, App	c 939	19.8	0.4	84	1	US-07-731-157A-33	Sequence 33, App
867	19.8	0.4	57	3	US-08-621-700-16	Sequence 16, App	c 940	19.8	0.4	84	1	US-08-208-886C-22	Sequence 22, App
868	19.8	0.4	57	3	US-08-929-940-16	Sequence 16, App	c 941	19.8	0.4	84	1	US-08-208-886C-34	Sequence 34, App
869	19.8	0.4	57	5	PCT-US95-03940-16	Sequence 16, App	c 942	19.8	0.4	84	1	US-08-053-131-125	Sequence 125, App
870	19.8	0.4	60	4	US-09-253-025-10	Sequence 10, App	c 943	19.8	0.4	84	1	US-08-704-744-32	Sequence 32, App
871	19.8	0.4	60	5	PCT-US93-01901-22	Sequence 22, App	c 944	19.8	0.4	84	1	US-08-704-744-34	Sequence 34, App
872	19.8	0.4	61	2	US-08-472-171-54	Sequence 34, App	c 945	19.8	0.4	84	1	US-08-645-641-125	Sequence 125, App
873	19.8	0.4	61	2	US-08-472-171-51	Sequence 34, App	c 946	19.8	0.4	84	1	US-08-469-557-22	Sequence 22, App
874	19.8	0.4	61	2	US-08-483-528B-50	Sequence 50, App	c 947	19.8	0.4	84	1	US-08-469-557-22	Sequence 22, App
875	19.8	0.4	61	2	US-08-894-526-34	Sequence 34, App	c 948	19.8	0.4	84	1	US-07-853-408B-125	Sequence 125, App
876	19.8	0.4	61	2	US-08-894-526-34	Sequence 34, App	c 949	19.8	0.4	84	1	US-08-096-762-125	Sequence 125, App
877	19.8	0.4	61	2	US-09-013-047-34	Sequence 34, App	c 950	19.8	0.4	84	2	US-08-290-793B-34	Sequence 34, App
878	19.8	0.4	61	2	US-09-013-047-34	Sequence 34, App	c 951	19.8	0.4	84	2	US-08-290-793B-34	Sequence 34, App
879	19.8	0.4	61	3	US-08-673-799C-50	Sequence 50, App	c 952	19.8	0.4	84	2	US-08-308-865-125	Sequence 125, App
880	19.8	0.4	61	3	US-09-374-597-51	Sequence 51, App	c 953	19.8	0.4	84	2	US-08-541-780-33	Sequence 33, App
881	19.8	0.4	61	3	US-09-374-597-51	Sequence 51, App	c 954	19.8	0.4	84	5	PCT-US92-10983-125	Sequence 125, App
882	19.8	0.4	61	4	US-09-393-385B-50	Sequence 50, App	c 955	19.8	0.4	85	1	US-08-627-497-4	Sequence 4, App
883	19.8	0.4	62	1	US-07-593-657-15	Sequence 15, App	c 956	19.8	0.4	85	4	US-09-060-410-11	Sequence 11, App
884	19.8	0.4	63	1	US-07-842-089E-5	Sequence 5, App	c 957	19.8	0.4	85	4	US-09-242-095-6	Sequence 6, App
885	19.8	0.4	63	1	US-08-264-485-5	Sequence 5, App	c 958	19.8	0.4	87	1	US-08-053-131-122	Sequence 122, App
886	19.8	0.4	63	2	US-09-003-081-7	Sequence 7, App	c 959	19.8	0.4	87	1	US-08-645-641-122	Sequence 122, App
887	19.8	0.4	63	3	US-08-648-262-7	Sequence 7, App	c 960	19.8	0.4	87	1	US-07-853-408B-122	Sequence 122, App
888	19.8	0.4	63	3	US-08-648-262-7	Sequence 7, App	c 961	19.8	0.4	87	1	US-08-096-762-122	Sequence 122, App
889	19.8	0.4	64	3	US-08-463-903-53	Sequence 53, App	c 962	19.8	0.4	87	2	US-08-308-865-122	Sequence 122, App
890	19.8	0.4	64	3	US-07-935-695-53	Sequence 53, App	c 963	19.8	0.4	87	4	US-09-208-966-27	Sequence 27, App
891	19.8	0.4	65	1	US-08-484-668B-34	Sequence 34, App	c 964	19.8	0.4	87	4	US-09-042-353-239	Sequence 239, App
892	19.8	0.4	65	1	US-08-484-668B-37	Sequence 37, App	c 965	19.8	0.4	87	4	US-08-758-417A-87	Sequence 87, App
893	19.8	0.4	65	3	US-08-718-904-91	Sequence 91, App	c 966	19.8	0.4	87	5	PCT-US92-10983-122	Sequence 122, App
894	19.8	0.4	65	3	US-08-463-160B-37	Sequence 37, App	c 967	19.8	0.4	88	1	US-07-964-624C-30	Sequence 30, App
895	19.8	0.4	65	4	US-08-463-160B-37	Sequence 37, App	c 968	19.8	0.4	88	1	US-08-442-062-30	Sequence 30, App
896	19.8	0.4	65	5	PCT-US91-02568-38	Sequence 38, App	c 969	19.8	0.4	88	1	US-08-748-697A-30	Sequence 30, App
897	19.8	0.4	66	1	US-07-627-539G-4	Sequence 4, App	c 970	19.8	0.4	88	4	US-09-165-616-30	Sequence 30, App
898	19.8	0.4	66	1	US-08-081-539-29	Sequence 29, App	c 971	19.8	0.4	88	4	US-09-171-759-17	Sequence 17, App
899	19.8	0.4	66	1	US-08-466-647-29	Sequence 29, App	c 972	19.8	0.4	88	4	US-09-143-014-7	Sequence 7, App
900	19.8	0.4	66	1	US-08-516-859A-96	Sequence 96, App	c 973	19.8	0.4	88	5	5194596-1	Patent No. 5194596
901	19.8	0.4	66	4	US-09-556-978B-42	Sequence 42, App	c 974	19.8	0.4	88	6	5219739-1	Patent No. 5219739
902	19.8	0.4	66	4	US-09-586-472-96	Sequence 96, App	c 975	19.8	0.4	89	4	US-09-020-956-63	Sequence 63, App
903	19.8	0.4	66	4	US-09-528-706-96	Sequence 96, App	c 976	19.8	0.4	89	4	US-09-030-607-63	Sequence 63, App

[illegible]

## pe 2 and

### Alaxia Type 2 and

Alaxia Type 2 and

EARLIER FILING DATE: 1996-07-18  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 13  
 LENGTH: 69  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 09-043-303-13

[illegible]

09-043-303-12  
 Sequence 12, Application US/09043303  
 Patent No. 6251589  
 GENERAL INFORMATION:  
 APPLICANT: TSUJI, SMOJI  
 APPLICANT: SANEI, Kazujiro  
 TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Ataxia Type 2 and  
 TITLE OF INVENTION: Primers Therefor  
 FILE REFERENCE: 0760-0241p  
 CURRENT APPLICATION NUMBER: US/09/043,303  
 CURRENT FILING DATE: 1998-05-18  
 EARLIER APPLICATION NUMBER: PCT/J96/01999  
 EARLIER FILING DATE: 1996-07-18  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 12  
 LENGTH: 78  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 09-043-303-12

**2586** CAACAGAGGGGACAGCAGCGCCAGCACCATATGCTGGGCTTACTGTAGACAGCGGCCGCCTC **2645**

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
**16** CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGCCGCCGCC **75**

LT 4  
8-240-0495-10  
quence 10, Application US/08240049B  
tent No. 5686239

GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.  
APPLICANT: Tam, Albert W.  
APPLICANT: Yarbough, Patricia O.  
TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods  
NUMBER OF SPOUNCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,049B

RESULT 5  
 US-08-542-634-12  
 Sequence 12, Application US/08542634  
 Patent No. 6214970  
 GENERAL INFORMATION:  
 APPLICANT: Fuerst, Thomas R.  
 APPLICANT: Mcatee, C. Patrick  
 APPLICANT: Yarbough, Patrice O.  
 APPLICANT: Zhang, Yifan  
 TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dehlinger & Associates  
 STREET: 350 Cambridge Ave., Suite 250  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94306  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/542,634  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fabian, Gary R.  
 REGISTRATION NUMBER: 33,875  
 REFERENCE/DOCKET NUMBER: 4600-0293.30  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 99 base pairs  
 TYPE: nucleic acid

```

1  APPLICANT: Fuerst, Thomas
2  APPLICANT: McAttee, Patrick
3  APPLICANT: Yarbough, Patrice
4  APPLICANT: Zhang, Yifan
5  TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods
6  NUMBER OF SEQUENCES: 24
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: Cathleen M. Desjardins, M.D.
9  STREET: 505 Penobscot Drive
10 CITY: Redwood City
11 STATE: CA
12 COUNTRY: US
13 ZIP: 94063
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Patentin Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/477,292
22 FILING DATE:
23 CLASSIFICATION:
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US/08/327,952
26 FILING DATE:
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Brookes, Allan A
29 REGISTRATION NUMBER: 36,373
30 REFERENCE/DOCKET NUMBER: G32P5
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (415)369-9500
33 TELEFAX: (415)368-0709
34 INFORMATION FOR SEQ ID NO: 12:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 99 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: double
39 TOPOLOGY: linear
40 MOLECULE TYPE: DNA (genomic)
41 HYPOTHEetical: NO
42 ORIGINAL SOURCE:
43 INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico strain)
44 INDIVIDUAL ISOLATE: 406.4-2 region
45
46 US-08-477-292-12

```

Patent No. 6277622

```
SOFTWARE: Patenlin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
```

Copied from PCI US2003092 on 04-03-2004



LENGTH: 84 base pairs

2670 2700

Copied from PCI-US2003092 on 04-03-2004







Copied from PCI US2003092 on 04-03-2004

ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/039,778B  
FILING DATE: 19930323  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/919,731  
FILING DATE: 27 July, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00537/068002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-039-778B-8

Query Match 0.5%; Score 24.2; DB 1; Length 69;  
Best Local Similarity 59.4%; Pred. No. 7.7e+03;  
Matches 41; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 2295 CCCATCATGACATCAACCATTTCCACATGCAACAGCCCTTCTTGCAGACGAGAG 2354  
DB 69 CCCCATTAAGCCCTGCTGATTCACAGCAGCCCTGCCCTGCTGATGCAAGACAGAG 10

QY 2355 AAACCCGCGC 2363  
DB 9 AACCTGCGC 1

RESULT 26  
US-08-240-049B-9  
Sequence 9, Application US/08240049B  
Patent No. 5686239  
GENERAL INFORMATION:  
APPLICANT: Reyes, Gregory R.  
APPLICANT: Tam, Albert W.  
APPLICANT: Yarbough, Patrice O.  
TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,049B  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Charles K. Sholtz  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)  
US-08-240-049B-9

Query Match 0.5%; Score 24.2; DB 1; Length 99;  
Best Local Similarity 53.8%; Pred. No. 8.8e+03;  
Matches 50; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3031 CAGCCGACGATGCTGCGCCACAGCGCTGAGAGGCGCCAGCCGCTGCGGACAGCGC 3090  
DB 7 CCGCCGACGACATCGGCTCGCTGCGCTGAGACAGCGCCAGCCGCGCTGCTGCTAC 66

QY 3091 TCCGAGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3123  
DB 67 GTCGTAGACCTACACAGCTGAGGCGCGCGCGCGC 99

RESULT 27  
US-08-542-634-11  
Sequence 11, Application US/08542634  
Patent No. 6214970  
GENERAL INFORMATION:  
APPLICANT: Fuerst, Thomas R.  
APPLICANT: McAtee, C. Patrick  
APPLICANT: Yarbough, Patrice O.  
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave., Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,634  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fadian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0293.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)  
INDIVIDUAL ISOLATE: 406.4-2 region  
US-08-542-634-11

Query Match 0.5%; Score 24.2; DB 4; Length 99;  
Best Local Similarity 53.8%; Pred. No. 8.8e+03;  
Matches 50; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 3031 CAGCCGACGATCGCTGGGCGGACGCGGTGAGAGGCGGACGCGGTCGGGACAGGC 3090  
DB 7 CCGCCGACGACGCTGCGCTGCGGTGACGAGGCCGCGCGCTTCCCTCAC 66

OY 3091 TCCGAGGCGCTGGCGCTGCTGTCGCGCGC 3123  
DB 67 GTGCTAGACCTACCAACAGCTGGGCGCGCGC 99

## RESULT 28

US-08-477-292-11  
Sequence 11, Application US/08477292

PATENT NO. 6291641  
GENERAL INFORMATION:  
APPLICANT: Fuerst, Thomas  
APPLICANT: McAtee, Patrick  
APPLICANT: Yarbrough, Patrice  
APPLICANT: Zhang, Yifan  
TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathleen M. Desjardins, M.D.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: US  
ZIP: 94063

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,292  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/327,952

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, Allan A  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: G32P5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)369-9500  
TELEFAX: (415)368-0709  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)  
INDIVIDUAL ISOLATE: 406.4-2 region  
US-08-477-292-11

Query Match 0.5%; Score 24.2; DB 4; Length 99;  
Best Local Similarity 53.8%; Pred. No. 8.8e+03;  
Matches 50; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 3031 CAGCCGACGATCGGCTGGGCGGACGCGGTGAGAGGCGGACGCGGTCGGGACAGGC 3090  
DB 7 CCGCCGACGACGCTGCGCTGCGGTGACGAGGCCGCGCGCTTCCCTCAC 66  
OY 3091 TCCGAGGCGCTGGCGCTGCTGTCGCGCGC 3123  
DB 67 GTGCTAGACCTACCAACAGCTGGGCGCGCGC 99

## RESULT 29

PCT-US95-13703-11  
Sequence 11, Application PC/TUS9513703  
GENERAL INFORMATION:  
APPLICANT:

TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13703  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0293.41  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0960  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)  
INDIVIDUAL ISOLATE: 406.4-2 region  
PCT-US95-13703-11

Query Match 0.5%; Score 24.2; DB 5; Length 99;  
Best Local Similarity 53.8%; Pred. No. 8.8e+03;  
Matches 50; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 3031 CAGCCGACGATCGGCTGGGCGGACGCGGTGAGAGGCGGACGCGGTCGGGACAGGC 3090  
DB 7 CCGCCGACGACGCTGCGCTGCGGTGACGAGGCCGCGCGCTTCCCTCAC 66  
OY 3091 TCCGAGGCGCTGGCGCTGCTGTCGCGCGC 3123  
DB 67 GTGCTAGACCTACCAACAGCTGGGCGCGCGC 99

## RESULT 30

US-08-053-131-129  
Sequence 129, Application US/08053131  
PATENT NO. 5661016  
GENERAL INFORMATION:  
APPLICANT: Lomborg, Nils



```

? APPLICANT: Kay, Robert M.
? TITLE OF INVENTION: Transgenic No. 5661016-Human Animals for
?   OF INVENTION: Producing Heterologous Antibodies
? NUMBER OF SEQUENCES: 197
? CORRESPONDENCE ADDRESS:
?   ADDRESSEE: Townsend and Townsend Kourile and Crew
?   STREET: One Market Plaza, Steuart Tower, Suite 200
?   CITY: San Francisco
?   STATE: California
?   COUNTRY: USA
?   ZIP: 94105
? COMPUTER READABLE FORM:
?   MEDIUM TYPE: Floppy disk
?   COMPUTER: IBM PC compatible
?   OPERATING SYSTEM: PC-DOS/MS-DOS
?   SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
?   APPLICATION NUMBER: US/08/053,131
?   FILING DATE: 26-APR-1993
?   CLASSIFICATION: 800
?   PRIOR APPLICATION DATA:
?     APPLICATION NUMBER: US 07/990,860
?     FILING DATE: 16-DEC-1992
?     PRIOR APPLICATION DATA:
?       APPLICATION NUMBER: US 07/810,279
?       FILING DATE: 17-DEC-1991
?       APPLICATION NUMBER: US 07/853,408
?       FILING DATE: 18-MAR-1992
?       ATTORNEY/AGENT INFORMATION:
?         NAME: Smith, William M.
?         REGISTRATION NUMBER: 30,223
?         REFERENCE/DOCKET NUMBER: 14643-9-3
?         TELECOMMUNICATION INFORMATION:
?           TELEPHONE: 415-326-2400
?           TELEFAX: 415-326-2422
?         INFORMATION FOR SEQ ID NO: 129:
?           SEQUENCE CHARACTERISTICS:
?             LENGTH: 69 base pairs
?             TYPE: nucleic acid
?             STRANDEDNESS: single
?             TOPOLOGY: linear
?             MOLECULE TYPE: DNA (genomic)
?             US-08-053-131-129
?
? Query Match          0.5%; Score 24; DB 1; Length 69;
? Best Local Similarity 68.8%; Pred. No. 8.5e+03;
? Matches 33; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

```

OY 4004 CAGTGGACAGGGGTACTGCTCACCACGCTCTCGGCGACAGCATGC 4051
|| |||| | ||| ||| | |||| | |||| | ||| |
DB 22 CACTGGGCGCAGGGACCCCTGCTCACCCTCTCTCAGGAGTCATCC 69

```

```

? RESULT 31
? US-08-645-641-129
? Sequence 129, Application US/08645641
? Patent No. 5719032
? GENERAL INFORMATION:
?   APPLICANT: Lomborg, Nils
?   APPLICANT: Kay, Robert M.
?   TITLE OF INVENTION: Transgenic No. 5719032-Human Animals for
?   OF INVENTION: Producing Heterologous Antibodies
? NUMBER OF SEQUENCES: 150
? CORRESPONDENCE ADDRESS:
?   ADDRESSEE: William M. Smith
?   STREET: Two Embarcadero Center, 8th Floor
?   CITY: San Francisco
?   STATE: California
?   COUNTRY: USA
?   ZIP: 94111-3834
? COMPUTER READABLE FORM:
?   MEDIUM TYPE: Floppy disk

```

```

? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
?   APPLICATION NUMBER: US/08/645,641
?   FILING DATE: 20-MAY-1996
?   CLASSIFICATION: 800
?   PRIOR APPLICATION DATA:
?     APPLICATION NUMBER: US 07/904,068
?     FILING DATE: 23-JUN-1992
?     ATTORNEY/AGENT INFORMATION:
?       NAME: Smith, William M.
?       REGISTRATION NUMBER: 30,223
?       REFERENCE/DOCKET NUMBER: 14643-000913
?       TELECOMMUNICATION INFORMATION:
?         TELEPHONE: 415-326-2400
?         TELEFAX: 415-326-2422
?       INFORMATION FOR SEQ ID NO: 129:
?         SEQUENCE CHARACTERISTICS:
?           LENGTH: 69 base pairs
?           TYPE: nucleic acid
?           STRANDEDNESS: single
?           TOPOLOGY: linear
?           MOLECULE TYPE: DNA (genomic)
?           US-08-645-641-129
?
? Query Match          0.5%; Score 24; DB 1; Length 69;
? Best Local Similarity 68.8%; Pred. No. 8.5e+03;
? Matches 33; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

```

OY 4004 CAGTGGACAGGGGTACTGCTCACCACGCTCTCGGCGACAGCATGC 4051
|| |||| | ||| ||| | |||| | |||| | ||| |
DB 22 CACTGGGCGCAGGGACCCCTGCTCACCCTCTCTCAGGAGTCATCC 69

```

```

? RESULT 32
? US-07-853-408B-129
? Sequence 129, Application US/07853408B
? Patent No. 5789650
? GENERAL INFORMATION:
?   APPLICANT: Lomborg, Nils
?   APPLICANT: Kay, Robert M.
?   TITLE OF INVENTION: Transgenic No. 5789650-Human Animals for
?   OF INVENTION: Producing Heterologous Antibodies
? NUMBER OF SEQUENCES: 150
? CORRESPONDENCE ADDRESS:
?   ADDRESSEE: William M. Smith
?   STREET: One Market Plaza, Steuart Tower, Suite 2000
?   CITY: San Francisco
?   STATE: California
?   COUNTRY: USA
?   ZIP: 94105
? COMPUTER READABLE FORM:
?   MEDIUM TYPE: Floppy disk
?   COMPUTER: IBM PC compatible
?   OPERATING SYSTEM: PC-DOS/MS-DOS
?   SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
?   APPLICATION NUMBER: US/07/853,408B
?   FILING DATE: 19920318
?   CLASSIFICATION: 800
?   ATTORNEY/AGENT INFORMATION:
?     NAME: Smith, William M.
?     REGISTRATION NUMBER: 30,223
?     REFERENCE/DOCKET NUMBER: 14643-9
?     TELECOMMUNICATION INFORMATION:
?       TELEPHONE: 415-326-2400
?       TELEFAX: 415-326-2422
?     INFORMATION FOR SEQ ID NO: 129:
?       SEQUENCE CHARACTERISTICS:
?         LENGTH: 69 base pairs
?         TYPE: NUCLEIC ACID
?         STRANDEDNESS: single

```

```

Oy      4004 CAGTGGGACAGGGGTACCCTGCCTCACACACTCTCTCGGGCAGACAGCATCC 4051
        ||||||| | ||||| | ||||||| ||||||| || ||||| |
Db      22 CACTGGGGCCAGGGCACCCCTGTCACCGCTCTCTCCTCAGGGAGTGATCC 69

RESULT 34
US-08-308-865-129
; Sequence 129, Application US/08308865
; Patent No. 5877397
; GENERAL INFORMATION:
; APPLICANT: Lomborg, Nils
; APPLICANT: Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 5877397-Human Animals for
; TITLE OF INVENTION: Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 150
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/308,865
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,707
; FILING DATE:
; APPLICATION NUMBER: US 07/904,068
; FILING DATE: 23-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14643-9-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-308-865-129

Query Match          0.5%; Score 24; DB 2; Length 69;
Best Local Similarity 68.8%; Pred.No. 8.5e+03;
Matches 33; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy      4004 CAGTGGGACAGGGGTACCCTGCCTCACACACTCTCTCGGGCAGACAGCATCC 4051
        ||||||| | ||||| | ||||||| ||||||| || ||||| |
Db      22 CACTGGGGCCAGGGCACCCCTGTCACCGCTCTCTCCTCAGGGAGTGATCC 69

RESULT 35
PCT-US92-10983-129
; Sequence 129, Application PC/TUS9210983
; GENERAL INFORMATION:
; APPLICANT: Lomborg, Nils
; APPLICANT: Kay, Robert M.
; TITLE OF INVENTION: Transgenic Non-Human Animals for
; TITLE OF INVENTION: Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith

```

```

FILING DATE: 23-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/16433
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,417
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/21803
FILING DATE: 01-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 014643-009040US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 249:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-042-353-249

Query Match 0.5%; Score 24; DB 4; Length 72;
Best Local Similarity 68.8%; Pred.No. 8.7e+03;
Matches 33; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 4004 CAGGGGACAGGGGTACTGGCTACACAGCAGCCTGGGCGACAGCATGC 4051
|| ||||| || ||||| || ||||| ||||| || |||||
Db 25 CACTGGGGCCAGGGCACCTGGTCACTGCTCAGGAGATGCATCC 72

RESULT 37
US-08-758-417A-97
; Sequence 97, Application US/08758417A
; Patent No. 6300129
; GENERAL INFORMATION:
; APPLICANT: Lomborg, Nils
; Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for

```

```

? NUMBER OF SEQUENCES: 417 Producing Heterologous Antibodies
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Townsend and Townsend and Crew LLP
? STREET: Two Embarcadero Center, Eighth Floor
? CITY: San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94111-3834
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US 08/758,417A
? FILING DATE: 02-DEC-1996
? CLASSIFICATION: <unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/728,463
? FILING DATE: 10-OCT-1996
? APPLICATION NUMBER: US 08/544,404
? FILING DATE: 10-OCT-1995
? APPLICATION NUMBER: US 08/352,322
? FILING DATE: 07-DEC-1994
? APPLICATION NUMBER: US 08/209,741
? FILING DATE: 09-MAR-1994
? APPLICATION NUMBER: US 08/165,699
? FILING DATE: 10-DEC-1993
? APPLICATION NUMBER: US 08/161,739
? FILING DATE: 03-DEC-1993
? APPLICATION NUMBER: US 08/155,301
? FILING DATE: 18-NOV-1993
? APPLICATION NUMBER: US 08/096,762
? FILING DATE: 22-JUL-1993
? APPLICATION NUMBER: US 08/053,131
? FILING DATE: 26-APR-1993
? APPLICATION NUMBER: US 07/990,860
? FILING DATE: 16-DEC-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Serafini, Andrew T.
? REGISTRATION NUMBER: 41,303
? REFERENCE/DOCKET NUMBER: 014643-009030US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 576-0200
? TELEFAX: (415) 576-0300
? INFORMATION FOR SEQ ID NO: 97:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 72 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA
? SEQUENCE DESCRIPTION: SEQ ID NO: 97:
? US-08-758-417A-97
?
Query Match 0.5%; Score 24; DB 4; Length 72;
Best Local Similarity 68.8%; Predt No. 8.7e+03;
Matches 33; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 4004 CAGTGGGACAGGCGTACTGTGCCTCCACCAAGCCTTCCTGCGCACAGCATGC 4051
|| ||| | | || | | | ||| | ||| | | | ||| |
Db 25 CACTGGGCCAGAGCCCTGTGTACCGCTCTCTCAGGAGATGCATCC 72

```

```

1  APPLICANT: Roninson, Igor B
2  TITLE OF INVENTION: BCL2 Derived Genetic Elements Associated
3  TITLE OF INVENTION: with Sensitivity to Chemotherapeutic Drugs
4  NUMBER OF SEQUENCES: 16
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE: Banner & Witcoff, Ltd.
7  STREET: 10 South Wacker Drive, Suite 3000
8  CITY: Chicago
9  STATE: Illinois
10 COUNTRY: USA
11 ZIP: 60606
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Patentin Release #1.0, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/405,702A
18 FILING DATE: 17-MAR-1995
19 CLASSIFICATION: 514
20 ATTORNEY/AGENT INFORMATION:
21 NAME: No. 5789389nan, Kevin E
22 REGISTRATION NUMBER: 35,303
23 REFERENCE/DOCKET NUMBER: 95,332
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: 312-715-1000
26 TELEFAX: 312-715-1234
27 TELEX: 910-221-5317
28 INFORMATION FOR SEQ ID NO: 5:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 84 base pairs
31 TYPE: nucleic acid
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: CDNA
35 US-08-405-702A-5
36
37 Query Match 0.5%, Score 24; DB 1; Length 84;
38 Best Local Similarity 58.3%, Pred. No. 9.2e+03;
39 Matches 42; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
40
41 QY 2632 GGCCTACCTGAGCAGCCGCCGCTCCTCAGAGGATCTCGCCCTTCCTCTCAGAGCCGCCGCTC 2681
42 3 GCTCCATCGAGCCCTCCGCGCAGCCGCGGAGCAGCATCTTCGCCGCCCTACCGCCGCACTT 62
43 Db 63 CGCGCAGATGTC 74
44
45 QY 2682 CAGCAGAGCGCTC 2693
46 1 1111 111
47 Db 63 CGCGCAGATGTC 74
48
49 RESULT 39
50 US-08-405-702A-9
51 Sequence 9, Application US/08405702A
52 Patent No. 5789389
53 GENERAL INFORMATION:
54 APPLICANT: Tarasewicz, Dariusz G
55 APPLICANT: Scholt, Brigitte
56 APPLICANT: Holzmayer, Tatiana A.
57 APPLICANT: Roninson, Igor B
58 TITLE OF INVENTION: BCL2 Derived Genetic Elements Associated
59 TITLE OF INVENTION: with Sensitivity to Chemotherapeutic Drugs
60 NUMBER OF SEQUENCES: 16
61 CORRESPONDENCE ADDRESS:
62 ADDRESSEE: Banner & Witcoff, Ltd.
63 STREET: 10 South Wacker Drive, Suite 3000
64 CITY: Chicago
65 STATE: Illinois
66 COUNTRY: USA
67 ZIP: 60606
68 COMPUTER READABLE FORM:
69 MEDIUM TYPE: Floppy disk
70 COMPUTER: IBM PC compatible
71 OPERATING SYSTEM: PC-DOS/MS-DOS
72

```

RESULT 40  
US-08-476-176B-22/C  
Sequence 22 Application US/08476176B  
Patent No. 5958708  
GENERAL INFORMATION:  
APPLICANT: Hardman, No. 5958708man  
APPLICANT: Kolbinger, Frank  
APPLICANT: Saldanha, Jose  
TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
immunoglobulin isotype  
TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
immunoglobulin isotype  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: No. 5958708artis Patent Department  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07936-1080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,176B  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/127,721  
FILING DATE: 27-SEPTEMBER-1993  
APPLICATION NUMBER: US 07/952,802  
FILING DATE: 25-SEPTEMBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5958708ak, Henry P.  
REGISTRATION NUMBER: 33,200  
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110

[illegible]

Search completed: December 11, 2002, 06:15:36  
Job time : 140 secs

